Topology Optimized Components for Mode- and Wavelength Division Multiplexing

Frellsen, Louise Floor

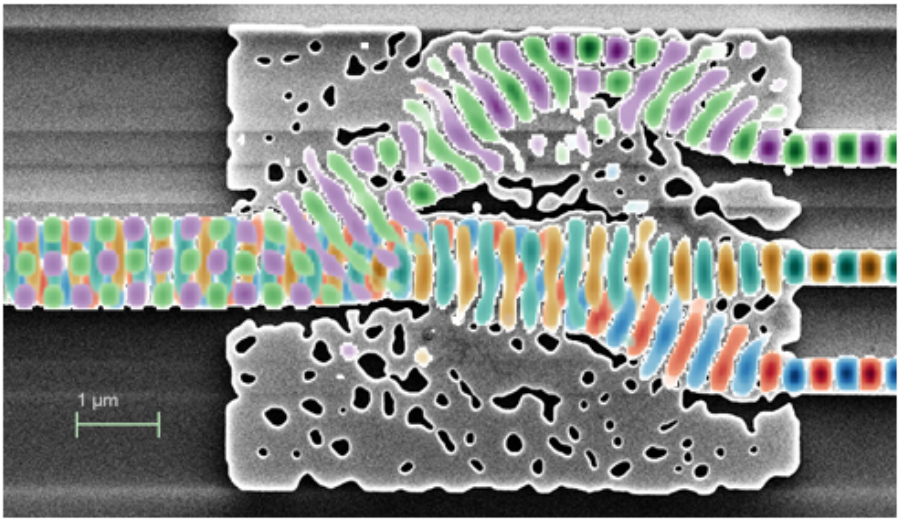
Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

Topology Optimized Components for Mode- and Wavelength Division Multiplexing

Ph.d. Thesis. DTU Fotonik - Department of Photonics Engineering

Louise Floor Frellsen



Supervisor Lars Hagedorn Frandsen
Co-supervisor Yunhong Ding

30th of September 2016
Technical University of Denmark

Abstract

This thesis deals with the topic of passive integrated nanophotonic devices realized in silicon on insulator material. The project has been concerned with all the steps of the process: Design, fabrication and characterization.

The focus has been on using the inverse design method topology optimization. The method has been employed extensively to obtain various nanophotonic components. One of the outcomes of the project has been the development of a solid understanding of the benefits and limitations of topology optimization.

This thesis presents numerous designs that have been characterized both through simulations and experiments. Among these are converters and (de-)multiplexers for mode division multiplexing, both realized with a record small footprint. Wavelength multiplexing devices were used as a basis for investigating the correlation between structure sizes and performance. Fortunately a larger footprint does not always give rise to better performance, however allowing for smaller feature sizes will. The design of compact tapers was commenced. Difficulties were met when working with very wide waveguides but methods for overcoming these were suggested.

A novel form of cladding modulated Bragg gratings, utilizing continuous rails to modify the refractive index and cause the reflections, has also been proposed and experimentally verified as part of this project.

This work has contributed additional components to the toolbox of devices necessary for integrated photonics. It has been shown that topology optimization is a strong method for creating extremely compact devices, the small features do however mean that they are not yet possible to fabricate on a large scale. Complex device functionalities can be obtained. Building on previous work of simpler structures it is comparatively easy to remake the new designs and then increase complexity without much impact on the footprint.

The benefit of inverse design tools, like topology optimization, is that they lead to structures without geometrical constraints and which are independent of the designer. This project has however shown, that the best results are obtained when iterating on the optimized structures and providing the tool with well-chosen starting point structures.

Resumé

Denne afhandling omhandler passive, integrerede nanofotoniske komponenter. Projektet har involveret alle processens faser: design fabrikation og karakterisering.

Fokus i projektet har ligget på brugen af den inverse designmetode topologioptimering. Metoden har været anvendt til at fremstille adskillige nanofotoniske komponenter. Yderligere har projektet bidraget til en udvidet forståelse af topologioptimerings fordele så vel som begrænsninger.

Denne afhandling indeholder talrige designs som er blevet karakteriseret gennem simuleringer så vel som eksperimenter. Iblandt disse komponenter er modekonvertere og (de-)multipleksere beregnet til mododelt multipleksing. Begge komponenttyper er opnået på rekordsmå arealer. Med afsæt i bølgelængdevis multipleksing er strukturernes funktionalitet undersøgt som en funktion af hele designets størrelse så vel som detaljegraden. Det blev fastslået, at det ikke altid er fordelagtigt at øge designområdets størrelse, imidlertid leder finere detaljer generelt til bedre funktionalitet. Arbejdet med at designe en kompakt forbredning af en bølgeleder blev påbegyndt. Imidlertid opstod komplikation i forhold til meget brede transmissionslinjer, men potentielle løsningsforslag er fremlagt.

En ny afart af kappemodulerede Bragg gitre er blevet introduceret. Her udnyttes kontinuerte rælinger med periodiske udsving til at forårsage de fornødne ændringer i brydningsindekset, som skal til for at skabe en refleksion. Funktionaliteten af disse strukturer er blevet eksperimentelt bekræftet.

Dette værk har således biddraget med en række nye komponenter, der kan føjes til de værktøjer, som er nødvendige for at udvikle brugen af integrerede fotoniske systemer. Det er blevet demonstreret, at topologioptimering er et effektivt værktøj til at forme enormt kompakte strukturer. De små detaljer i designsne gør dog, at det endnu ikke er muligt at masseproducere dem. Ved først at designe relativt simple strukturer er det forholdsvist nemt at udvide kompleksiteten af anordningerne, og dette kan gøres uden stor øgning af de enkelte strukturers areal.

Fordelen ved inverse designmetoder, så som topologioptimering, er at de leder til strukturer uden geometriske begrænsninger og som i sig selv er uafhængige af designeren. Dette projekt har imidlertid vist, at de bedste resultater opnås ved at gentage optimeringerne med afset i tidligere versioner, så programmet har gode start-geometri at arbejde ud fra.

Acknowledgements

First and foremost I would like to thank my competent main supervisor Senior Researcher Lars Hagedorn Frandsen and co-supervisor Researcher Yunhong Ding. They taught me much about nanophotonics, research, handling challenges on my own and asking for help when needed. Without their guidance I would never have navigated my way through this project.

There are many others who I owe thanks for having helped me through the project too.

All of the nanophotonic devices group, led by Professor Kresten Yvind, has contributed to a productive and comfortable work environment necessary for any sort of good work to take place. Especially I would like to thank the two members I have worked most closely with. Miranda Mitrovic has freely shared her insight and my frustrations when needed. Xiaowei Guan has always kindly offered help and guidance and more than once his knowledge on the cleanroom processes has saved me days of delay.

I was fortunate enough to have not just one but two groups during my project. A huge thanks goes to all of the great people of the Silicon Integrated Nanopotonics group at IBM who took care of me during my time in New York. Especially I would like to mention my two local supervisors Jason S. Orcutt and Jessie C. Rosenberg who included me, guided me and made sure that I learned and produced during my stay. Also a warm thank you to group leader William M. J. Green, although we did not work together much during my stay, it was clearly important to him that I had a good experience while there.

A big thanks to Professor Ole Sigmund from DTU Mekanik for sparring and giving advice on the more technical side of topology optimization. It was exciting to see just how much this method can do. Had it not been for the time limit of the project, I would have enjoyed to see how these additional tweaks would work in the field of nanophotonics. Also a special thanks to Tine Greibe from Danchip who was extremely patient in coming to my rescue every time I had trouble with the e-beam writer. She did not only fix the problems but also taught me more about how the machine worked every time. I am also grateful for the pleasant company of my office mate Krzysztof Iwaszczuk who was always a great help in figuring out all of the framework from the inner workings of DTU Fotonik to how to handle publications and conferences. In that context I would also like to thank the helpful secretaries of DTU Fotonik who made sure that everything runs smoothly at the department, and who always seem to know all the answers.

This work was financed by the VILLUM foundation via the project "Optical Nano-engineered Components for High-Capacity Integrated silicon Photonics" (ONCHIP). Furthermore I am grateful for the generous financial aid in covering the expenses of my external stay that I received from "Ulla og Mogens Folmer Andersens Fond" and "Otto Mønsted Fond".

Finally a special thanks goes to my family and friends without whom I would never have had the energy to see this through. They supported me when I worked, told me when I needed to take a break, and without them there would be many more typos in this thesis.

List of publications

L. H. Frandsen, Y. Elesin, L. F. Frellsen, M. Mitrovic, Y. Ding, O. Sigmund, and K. Yvind, "Topology optimized mode conversion in a photonic crystal waveguide fabricated in silicon-on-insulator material", *Optics Express*, vol. 22, pp. 8525-8532, 2014.

L. F. Frellsen, L. H. Frandsen, Y. Ding, Y. Elesin, O. Sigmund and K. Yvind, "Topology optimized silicon photonic wire mode-multiplexer", *Photonics West Conference Proceedings*, 2015.

L. F. Frellsen, L. H. Frandsen, Y. Ding, Y. Elesin, O. Sigmund and K. Yvind, "Topology optimized design for silicon-on-insulator mode converter", *IEEE Photonics Conference*, pp. 162-163, 2015.

Y. Ding, L. F. Frellsen, X. Guan, J. Xu, F. da Ros, H. Ou, C. Peucheret, L. H. Frandsen, L. K. Oxenløve and K. Yvind, "On-chip mode division multiplexing technologies", *Proceedings of the Spie*, vol. 9774, 2016.

L. F. Frellsen, Y. Ding, O. Sigmund, and L. H. Frandsen, "Topology-optimized mode converter in a silicon-on-insulator photonic wire waveguide", *Conference on Lasers and Electro-Optics, OSA Technical Digest*, paper STh3E.4, 2016.

L. F. Frellsen, Y. Ding, O. Sigmund, and L. H. Frandsen, "Topology optimized mode multiplexing in silicon-on-insulator photonic wire waveguides", *Optics Express*, vol. 24, pp. 16866-16873, 2016.

L. F. Frellsen, Y. Ding, O. Sigmund and L. H. Frandsen, "Topology optimized design of a transverse electric higher order mode converter", *IEEE Photonics Conference*, 2016.

List of abbreviations

2MUX - 2 mode (de-)multiplexer, specifically operating on the TE_0 and TE_1 modes.

3MUX - 3 mode (de-)multiplexer, operating on the TE_0, TE_1 and TE_2 modes.

C-DRS - Cladding-modulated distributed resonant structure.

DBR - Distributed Bragg reflector.

DUV - Deep ultra violet, referring to the wavelength range around 250 nm.

EBL - Electron beam lithography, also using the abbreviation of e-beam.

FD - Finite difference.

FDTD - Finite difference time domain.

FSR - Free spectral range.

IPA - isopropyl alcohol.

IR - Infra red.

MDM - Mode division multiplexing.

PEO - Point energy objective.

PER - peak extinction ratio.

PDE - Partial differential equation.

PhC - Photonic crystal.

PhW - Photonic wire, being an on-chip waveguide.

PIC - Photonic integrated circuit.

PSO - Point shape objective.

RIE - Reactive ion etching.

SDM - Space division multiplexing.

S-DRS - Sidewall-modulated distributed resonant structure.

SEM - Scanning electron microscopy.

Si - Silicon.

SiO_2 - Silica.

SOI - Silicon on insulator, in this project being silicon on silica.

TM - Transverse electric.

TM - Transverse magnetic.

TO - Topology optimization.

WDM - Wavelength division multiplexing.

Contents

Contents **xi**

1 Introduction **1**
1.1 Optical communication . . . 1
1.2 Integrated silicon photonics . . . 2
1.3 This project - topology optimization for integrated silicon phtonics . . . 4

2 Topology Optimization **7**
2.1 The method of topology optimization . . . 9
2.1.1 Tasks for the designer . . . 11
2.2 Other optimization methods . . . 17

3 Fabrication of Passive Nanophotnic Structures **19**
3.1 Process Overview . . . 19
3.2 Electron beam Lithography . . . 22
3.2.1 Sample design . . . 23
3.3 Reactive Ion Etching . . . 24

4 Passive Characterization Methods **25**
4.1 Mode profile recording . . . 26
4.2 Insertion loss measurements . . . 27
4.3 Grating measurements . . . 29

5 Mode division multiplexing components **31**
5.1 Introduction . . . 31
5.2 Mode Conversion . . . 32
5.2.1 Photonic crystal mode converter . . . 33
5.2.2 Straight Mode Converter . . . 35
5.2.3 S-bend structure . . . 37
5.2.4 Optimized bend mode converter . . . 39
5.2.5 Re-optimization of the mode converter . . . 42
5.2.6 Higher order mode converter . . . 44
5.3 2-mode Multiplexer . . . 47
5.3.1 Initial design attempts . . . 50
5.3.2 Square input structure . . . 55
5.4 3-mode Multiplexer . . . 59
5.5 Design Robustness . . . 64
5.5.1 Resolution and filters . . . 64
5.5.2 Structural changes . . . 67
5.5.3 Temperature variations . . . 70
5.6 Summary and outlook . . . 72

6 Size-performance-correlation in wavelength multiplexers **75**
6.1 Introduction . . . 75
6.2 Size-performance-correlation . . . 77
6.2.1 Square structure . . . 79
6.2.2 Rectangular structure . . . 81
6.2.3 Asymmetric waveguides . . . 83
6.2.4 Waveguide spacing . . . 85

6.3 Feature size-performance-correlation . . . 86
6.3.1 Filter sizes . . . 86
6.3.2 Reapplying filters . . . 87
6.4 Comparison to other studies . . . 90
6.5 Summary and discussion . . . 92

7 Compact Taper **95**
7.1 The wide fundamental mode . . . 97
7.2 Optimized designs . . . 100
7.2.1 Tapered input . . . 100
7.2.2 2D investigation of taper length . . . 100
7.2.3 3D optimized tapers with swapped source and objective . . . 102
7.2.4 Alternative optimization attempts . . . 104
7.3 Small difference taper . . . 105
7.3.1 Stepwise taper design . . . 106
7.4 Summary and discussion . . . 108

8 Continuous cladding modulated Bragg gratings **109**
8.1 Introduction . . . 109
8.2 Step index gratings . . . 114
8.2.1 Step index simulations . . . 114
8.2.2 Step index measurements . . . 117
8.3 Continuous width grating . . . 120
8.3.1 Continuous width simulations . . . 120
8.3.2 Continuous width measurements . . . 124
8.4 Summary . . . 133

9 Future prospects and Conclusion **135**

References **139**

CHAPTER 1
INTRODUCTION

1.1 Optical communication

In the 25 years the internet has been available the traffic across it has been growing exponentially [1] as presented in figure 1.1.1. Presumably this was spurred on by the technological advances that continued to increase the capacity. Since it has become broadly available, it has become an everyday need for many people - and still more are getting connected. As we are getting the option to share more, we are also creating more things to share. Due to the huge increase in user generated data in the form of emails, blog posts and pictures, 90 % of all data has been generated within the last 2 years [2]. Putting an extra strain on the network, we have now also gotten used to receiving all of this data on demand, even in large amounts, with streaming services like Netflix and Spotify becoming ever more popular.

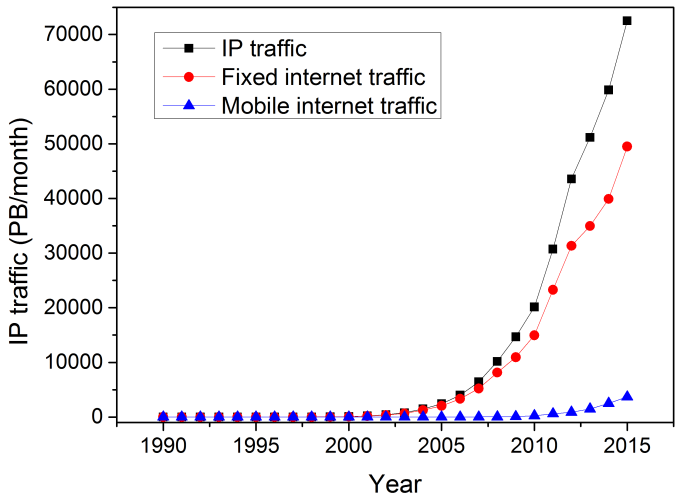


Figure 1.1.1: Average IP traffic per month development through the lifespan of the internet. Fixed internet traffic indicates connections which are wired while mobile is wireless traffic. The total IP traffic is the sum, plus additional internet backbone traffic. Data from Cisco [1].

One of the technologies that helped make this development possible was - and still is - optical fibres. In the last decades, optical communication technology has made enormous progress and many steps have been taken to increase the data carrying capacity of the fibres [3]. Among the already widely implemented methods are the

use of erbium doped fibre amplifiers and wavelength division multiplexing (WDM). For the latter the transmission is increased by sending different signals carrying diverse data in the same fibre simultaneously but allowing them to be distinguished by their different wavelengths. It has been suggested that the addition of multiplexing the signal in space will be the next method for increasing the capacity. There are multiple ways of doing this, the simplest begin to add more fibres, which although functional is not a very efficient solution. Having multiple cores in the same fibre is another way of getting around this, but the method of greatest interest to this thesis is the scheme of mode division multiplexing (MDM). Fibres can potentially support hundreds of modes but as of yet only single- and few-mode fibres are in use for optical communication systems. Although this project has not been concerned with the development of fibre technology, it is closely linked. There is no getting around that when working with photonic integrated circuits (PICs), they will in the long run need to be compatible with fibre optics. Hence it is important that any signal which may be coupled onto the chip can be treated directly on it. For this reason, as well as to have the highest transfer capacity on chip also, developing WDM and MDM for on-chip devices is of great interest.

1.2 Integrated silicon photonics

Photonic integrated circuits have received much attention lately because of their great potential for data communication across moderate and short ranges. The current technology employed on these ranges is based on copper wires. The copper technology has undergone huge development in the last decades. The technology is however nearing its physical limits [4] and alternative solutions with greater bandwidths are being sought after. Although it is likely that copper will always play an important role in the field of communication, the incorporation of PICs and a segmentation of the processing tasks offer a potential solution to its limitations [5]. The long term goal is to be able to incorporate optical components along with the electronic devices on the same chips. In that manner some of the optical signal could be coupled from the fibres directly onto the chip and processed there. Aside from adding capacity and overcoming the electronics interconnect bottleneck, this would also lead to energy savings in regards to the conversion of the optical signal to an electronic one. Additionally it has the potential of increasing processing speeds by allowing for a higher capacity on the same physical scale.

When it is the goal to interact with an already securely established industry, such as electronic communication, there are a number of things which are important to consider. For one, the new technology has to better or cheaper than the existing and it must be easily accessible as well. These issues of compatibility can be put into a number of categories as summed up in figure 1.2.1 and will be discussed in more detail here.

Material compatibility is important when the goal is to integrate a new technology with an existing one. There are a number of different materials that can allow for lattice constant matching, however for doped semiconductors the metal present in electronic systems cause a contamination problem. It is thus natural to turn to not just a compatible material, but the same material, namely silicon (Si). If doing so the material compatibility is no longer an issue, however it does of course require silicon to be a feasible solution for the optical nanophotonics. Luckily it is. Silicon has a number of material qualities that makes it very attractive for photonic devices. The high index contrast between silicon ($n_{Si} = 3.45$) and its oxidised form

Material compatibility	Process compatibility
- Lattice constants - Contamination - Bonding	- Fabrication - Foundry scale
Economic compatibility	**Heat compatibility**
- Area utilization - High-volume market	- Low heat dissipation - High temperatures - Refractive index - Thermal expansion - Propagation loss

Figure 1.2.1: Simple sketch of the four cornerstones of compatibility requirements for photonic integrated circuits along with keywords of relevant factors.

silica ($n_{SiO_2} = 1.45$) makes it possible to create very compact optical structures. The cross-sections of single mode waveguides can be scaled down to $< 400\,\mathrm{nm}^2$. This also allows for compact waveguide bends which will be necessary to route the signal efficiently on chip. At the same time, the large index contrast allows for high intensity signals that in turn makes nonlinear interactions possible. The benefit of the silicon platform does, however, also introduce challenges: Because of this high index contrasts, surface roughness will cause significant scattering and thus induce propagation losses. This puts high demands on the fabrication of the devices since low losses are, of course, a high priority in any optical device. Another challenge of the use of Si is the large birefringence. This will often be handled by working only with either the transverse electric (TE) or transverse magnetic (TM) polarization. In this project only the TE mode has been addressed. The direct material compatibility does also work in favour of the optical solution because of the availability of high quality silicon on insulator (SOI) substrates.

Process compatibility refers to the possibility to include the fabrication of the photonic circuits in the same process flow as the current electronic devices. This is important as it mitigates the need of a new industry being formed and because it will be possible to rely on existing supply chains. This point is partially solved when working with silicon as the material, because it has direct compatibility with the complimentary metal-oxide-semiconductor processes.

Economic compatibility is a large and complex challenge and it is closely linked with all of the other compatibilities. Consumers may be willing to pay for new technology, but in a case like this, where the incorporation of PICs will enable an increase in the capacity rather than introduce an entirely new functionality, an acceptable rise of the price will be low. The development of CMOS has been focused on scaling down feature lengths to most efficiently utilize the chip area. Hence it is essential that PICs will not significantly increase the size of the chips [6]. It is thus important that device designs have small footprints. Satisfying this subset of the economic compatibility aspect of PICs is the main focus of this thesis. Another challenge when it comes to the cost of PICs is the economy of scale. This means that the products must have a high volume market to fully use the potential of the process compatibility. In addition mass production allows the high engineering cost

to be distributed and shared across a higher number of devices. Unfortunately the former point on needing small devices complicates fulfilling the latter requirement. Very small optical structures can be fabricated well using electron beam (e-beam) lithography, as will be touched upon in chapter 3. This is, however, not a technology which can easily be scaled to mass production. The demand for large scale production creates additional requirements for the device design. It will be necessary to find an appropriate trade-off between footprint, feature sizes and fabricability.

Heat compatibility is the last of the compatibility issues and one for which an apparent solution is not currently present. When the temperature changes it will affect the refractive index of Si [7] in addition to causing expansion/contraction of the geometrical structures. Both of these effects may decrease the functionality of the integrated devices. Although various types of semiconductor lasers meant for Si substrates are being developed, currently PICs tend to rely on light sources off-chip. In regards to heat dissipation this is actually a benefit as it removes the strongest source of heating. However integrated functionalities, such as amplifiers, will still cause temperature increases on the chip and thus it is a challenge which needs to be addressed. Furthermore PICs integrated with electronics will also need to be able to handle their large power density which is on a much higher scale. This challenge has not been a focus of the present project, but is lightly touched upon in section 5.5.

The use of silicon for PICs does in itself fulfil two of the compatibility issues. It is the task of the device designers to ensure that the other two can be met as well. In this thesis, the focus has been on the aspect of scaling and size of the optical devices needed for the integration on chips.

1.3 This project - topology optimization for integrated silicon phtonics

The work described in this thesis was carried out during the three years of the Danish Ph.d. study. The majority of this time was spent at the Technical University of Denmark (DTU), department of photonics engineering. Time was split between the tasks of designing, fabricating and characterizing optical devices. The design process was done in collaboration with DTU's department of mechanical engineering and fabrication took place in the Danchip cleanroom facilities. Half a year of the project was spent at the IBM T. J. Watson Research Centre in Yorktown, New York, working in the nanophotonics group.

The guiding of light in PICs is done through the use of ridge waveguides, sometimes also called photonic wires (PhW). They use the concept of total internal reflection to confine light within a core material which has a higher refractive index than the cladding. As mentioned, the large difference in refractive index between silicon and potential claddings such as silica or air, lead to tight confinement of the optical modes. Simple tasks such as the guiding, bending and splitting of optical signals are easy to realize thanks the tight confinement of the light and the material's transparency at optical wavelengths for communication. It does however take more complex scattering and interference processes to create integrated devices with more diverse functionalities.

The goal of this project has been to explore the potential of passive silicon nanophotonic devices for the routing and processing of optical signals on chip. Full

system integration has however been outside the scope of this work. Rather it was the aim to develop a number of novel designs which will be added to the toolbox of optical signal processing devices. The majority of the work has been done applying the inverse design method topology optimization (TO) for the device design. Currently there are two main approaches to the design of such devices: classical parametric search designing and inverse design. These will be described and compared more in depth in chapter 2. The main benefit of the inverse design method, which TO belongs to, is that it is free of any inherent geometrical constraints. By creating odd scattering structures, that the designer would not have thought of, the performance of the devices may be increased significantly. An important factor for PICs is the footprint; creating devices that allows for efficient use of chip area has been a focus. This includes the development of multiplexing devices functioning on modes and wavelength, respectively. These devices allow for the use of less waveguides running across the chip when routing the signals. Minimizing the footprints of these devices has been a goal, and in doing so we have approached the fundamental lower limits of the devices sizes. To increase our understanding of both the method of inverse design and the size limitations of the scattering devices the correlation between device and feature sizes and the functionality was investigated. Understanding this relation is also of great relevance in regards to the possibility of fabricating such compact devices. Hence this work has helped map out the capabilities and limitations of TO in its current form. Opening up the software and implementing changes that can remove some of these limitations will be work for the future.

In addition to this, the classical parametric search design method has also been employed in this project during the collaboration with IBM. This was done as part of the design of a novel method of creating Bragg gratings to function as the reflectors in integrated lasers. Due to the periodicity and length of such devices, along with the goal of creating something which will be easy to fabricate, the classical design method was more suited.

The present thesis includes nine chapters, where this the first chapter is the introduction and the last is a conclusion and outlook of the project as a whole. **Chapter 2** describes, in more detail, the method of topology optimization, which has been used for the majority of the design tasks. **Chapter 3** and **chapter 4** deal with the physical frames for the project, describing the standardized processes used for fabrication and characterization of the devices, respectively. **Chapter 5-8** then presents the results produced during this project. **Chapter 5** describes the design of devices for mode conversion and multiplexing including a comparison between simulations and the experimental results from the fabricated devices. **Chapter 6** offers a simulation-based analysis of the correlation between device footprint/feature sizes and the performance. The analysis was conducted specifically on wavelength multiplexing devices. **Chapter 7** includes the preliminary work on developing compact tapers designed using TO. Finally **Chapter 8** presents the work done in collaboration with IBM including the numerical and experimental analysis of cladding modulated Bragg gratings.

CHAPTER 2

TOPOLOGY OPTIMIZATION

Topology optimization is an inverse design tool which is well-established in the field of mechanical engineering [8] and it has been an important industrial tool especially for weight optimization of machine parts [9]. In recent years it has been employed in other fields for optimizations of such things as micro-electro-mechanical systems [10], microfluidics [11] and since 2004 also for the designs of PICs [12].

For many of the possible PIC components, such as photonic crystals, surface plasmon structures, and meta materials the functionality arises from the placement and distribution of dielectric materials, metals, and air [13]. Topology optimization is a technique for determining the necessary placement. It is an inverse method, which is to say that a certain objective, such as the field intensity, is defined from the beginning and the tool will then find the material distribution which best obtains said objective for a given input. One of the great benefits of TO is that there are no inherent constraints on the geometries that can be used in the device; this allows for unintuitive structures rather than shapes that human beings would turn to - such as circles and squares. It should of course be noted that it can be beneficial to impose some form of constraints, through for example filtering, to ensure that the structures are possible to fabricate and to achieve robustness.

The basic concept of TO is to utilize finite element or finite difference (FD) analysis, in which a starting point structure is defined by a pixel, or voxel in 3D, representation (finite elements). Each pixel is given a value between 0 (air) and 1 (solid dielectric material). Repeated FD analysis is then made to evaluate the performance of the structure and the material is iteratively redistributed to improve it. A gradient based sensitivity analysis is employed to limit the material redistribution to areas of importance and to guide the device towards convergence, which is usually obtained after ~ 200 iterations. Towards the end filtering ensures that a minimum of grey areas exist, that is to say that the fewest possible pixels have a value in between 0 and 1. In figure 2.0.1 a conceptual overview of the procedure, illustrated by examples of a developing design, is shown.

Although a benefit of the inverse design method is how it allows for the optimum design for a specific functionality independent of a designer, there are still numerous parameters that must be chosen by a person. These include the determination of the area in which material redistribution can take place denoted the design domain (DD), the starting point structure in general, the choice and placement of source and objectives, as well as filtering demands.

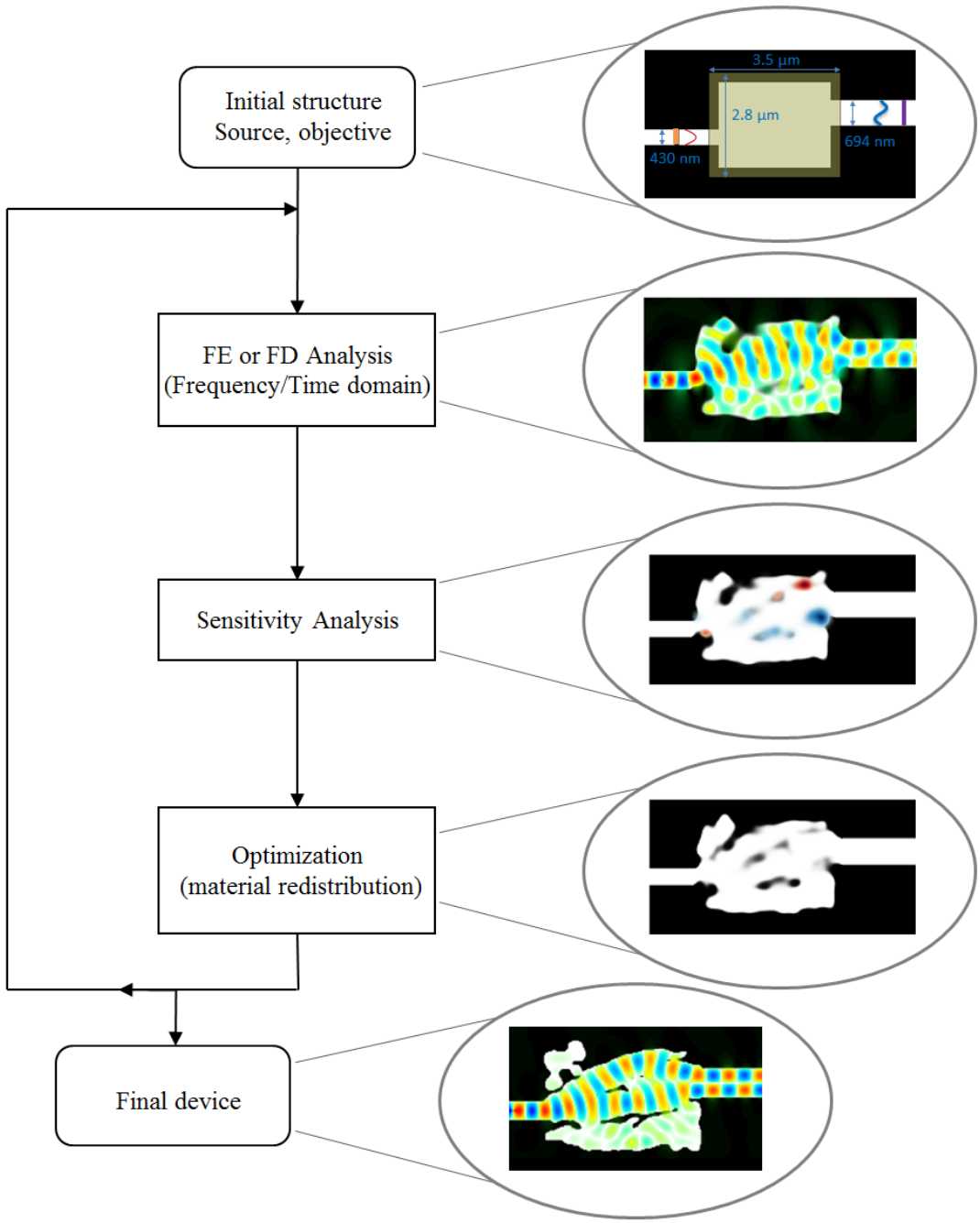


Figure 2.0.1: Schematic illustration of the concept of topology optimization. The squares in the center indicate a given step in the procedure while the ovals to the right contain corresponding design examples from the mode converter design presented in section 5.2. To the left the development of the same design is illustrated, to show what happens during the many iterations.

In this chapter the theory of topology optimization will be described cursorily with a focus on the aspects employed in the optimization software Phazor which has been developed in-house [14, 15]. Two different types of objectives can currently be used in this software and they will be described separately in this chapter.

2.1 The method of topology optimization

Phazor, the software used for the TO, has been developed previous to this project and although additional functionalities could be implemented and optimized doing so is outside the scope of the present work. What is presented here is thus an understanding of the method based on the literature (especially [13], [15] and [14]) and conversations with specialists, not personal experience.
The photonic device behaviour is governed by Maxwell's partial differential equations (PDEs)

$$\nabla \times \mathbf{H} = \epsilon_0 \epsilon_r \frac{\partial \mathbf{E}}{\partial t} \tag{1}$$

$$\nabla \times \mathbf{E} = -\mu_0 \frac{\partial \mathbf{H}}{\partial t} \tag{2}$$

where the electric and magnetic vector fields are $\mathbf{E} = (E_x, E_y, E_z)^T$ and $\mathbf{H} = (H_x, H_y, H_z)^T$. ϵ_0 and ϵ_r are the freespace and the relative permittivity, respectively while μ_0 is the freespace permeability. It is possible to simplify this to a 2D calculation by considering either TE or TM polarization, this reduced the calculation time noticeably, however in most cases 3D simulations were found to reveal better correspondence with measured results as will be expanded upon in later sections. The calculations can be done in the time domain using a numerical integration algorithm. Alternatively the calculations are performed in the frequency domain in order to find the steady state behaviour of the time-harmonic excitations expressed as:

$$E_z(x,t) = E_z(x)e^{i\omega t} \tag{3}$$

$$H_z(x,t) = H_z(x)e^{i\omega t} \tag{4}$$

The boundary conditions used to minimize reflection of waves, in spite of the many differing incident angles of light from the highly scattering devices, are perfectly matched layers (PMLs). These are layers of extra absorbing domains all around the computational domain and mathematically they include complex damping functions that will ensure that the waves are dissipated within. The strength of the dampening is of great importance. If it is too high reflections will occur at the interface to the PML while if it is too low reflections will still occur at the interface after the PML.

Material redistribution is essential to TO and for this to be possible it is necessary to introduce parametrization of the structure allowing the material properties of each local point to be optimized separately. As the optimization algorithm is gradient based it is necessary to introduce a continuous design variable:

$$0 \leq \rho \leq 1. \tag{5}$$

This variable can then be implemented in linear material interpolation laws. As an example, the law for the permittivity in the case of TE-polarized light may look as:

$$\frac{1}{\epsilon_r} = \frac{1}{\epsilon_{r1}} + \rho \left(\frac{1}{\epsilon_{r2}} - \frac{1}{\epsilon_{r1}} \right). \tag{6}$$

Unfortunately allowing the material variable to be continuous, as needed for the sensitivity analysis, leads to the existence of "grey" areas which are neither of the two materials. In some of the devices where there are many waves reflected locally, this problem is very little as the grey areas have a lower index contrast and thus decreases reflection making it inherently unfavourable. Otherwise penalization techniques can be introduced, for example by imposing artificial damping in grey areas.

Sensitivity analysis is performed using the adjoint approach; rather than having to perform independent optimizations for each location in the grid, an adjoint set of equations is made. It is assumed that it is possible to define an objective function, which is a measure of the performance of the device. This function will be dependent on the field vector $\mathbf{u}$ describing either the E- or H-field and the material parameter ρ, written in the frequency and time domain, respectively, as:

$$\Phi_{FD}\left(\mathbf{u},\rho\right) \text{ and } \Phi_{TD} = \int_0^T \phi\left(\mathbf{u},\rho\right)\,\mathrm{d}t \tag{7}$$

where T is the total simulation time for the transient computation. The sensitivity of the objective function is then found by deriving it with respect to all design variables. This is in practice made possible by applying the chain rule:

$$\frac{\mathrm{d}\Phi}{\mathrm{d}\rho} = \frac{\partial\Phi}{\partial\rho} + \frac{\partial\Phi}{\partial\mathbf{u}}\frac{\partial\mathbf{u}}{\partial\rho}. \tag{8}$$

However it is not straight forward to compute the implicit derivative $\frac{\partial\mathbf{u}}{\partial\rho}$, and this is where the adjoint approach enters. Taking the frequency domain as the simpler example the adjoint equation is

$$\mathbf{S}^T\lambda = -\frac{\partial\Phi_{FD}}{\partial\mathbf{u}} \tag{9}$$

where $\mathbf{S} = -\omega^2\mathbf{M} + i\omega\mathbf{C} + \mathbf{K}$ is the system matrix containing matrices formed during the discretization procedure and which is part of the linear equation set $\mathbf{Su} = \mathbf{f}$. Hence the final expression for the sensitivity is

$$\frac{\mathrm{d}\Phi_{FD}}{\mathrm{d}\rho} = \frac{\partial\Phi_{FD}}{\partial\rho} + 2\Re\left(\lambda^T\frac{\partial\mathbf{S}}{\partial\rho}\mathbf{u}\right). \tag{10}$$

The great benefit of this method is that the adjoint equation (9) needs only be solved once, regardless of the number of design variables. In this way the additional computational cost for each objective function is no more than one extra backward FDTD time integration. This is the same for the transient problem, which also requires only a single additional system calculation.

To avoid very small features in the devices, as these are not readily fabricated, it is possible to introduce size filters. This can be done in multiple ways either applying the filter to the sensitivity analysis or filtering the density of the material. In the current version of Phazor a Heaviside filter is applied. This is done by combining a hat-filter with a Heaviside-type projection operator [16]. For the hat-filter a window of a user-specified radius, R, is introduced. It defines a neighbourhood N_e of each

element e (that is a single voxel when working in 3D) as the elements lying within this radius

$$N_e = \{i | \|\mathbf{x}_i - \mathbf{x}_e\| \leq R\}, \tag{11}$$

where $\mathbf{x}$ is the positions vector of centroid of the finite elements. The hat-filter can be considered a mapping $\rho \mapsto \hat{\rho}$ regulating the densities in accordance with

$$\hat{\rho}_e = f_e^t(\rho) = \frac{\Sigma_{i \in N_e} w_i(\mathbf{x}_i)\, \rho_i}{\Sigma_{i \in N_e} w(\mathbf{x}_i)}. \tag{12}$$

Here w_i is the filter weight which is a linearly decaying function $w(\mathbf{x}_i) = R - \|\mathbf{x}_i - \mathbf{x}_e\|$. This part introduces the length scale of R into the filter, and the Heaviside is then a mapping $\rho \mapsto \tilde{\rho}$ including the two filter parameters η and β which can be used to more finely control the effect of the filter:

$$A_e = 1 - \frac{\hat{x}_e}{\eta}, \tag{13}$$

$$B_e = \frac{\hat{x}_e - eta}{1 - \eta} \tag{14}$$

$$\tilde{x}_e = f_e^h(\rho) = \begin{cases} \eta \left(e^{-\beta A_e} - A_e e^{-\beta}\right), & \text{if } 0 \leq \hat{\rho}_e \leq \eta \\ (1-\eta)\left(1 - e^{-\beta B_e} + B_e e^{-\beta}\right) + \eta & \text{if } \eta < \hat{\rho}_e \leq 1 \end{cases}. \tag{15}$$

Having introduced this length scale R it becomes the limiting factor for feature sizes or holes within the optimized structure. However this filtering will risk some additional grey areas. To minimize this effect the filter size is gradually decreased throughout the optimization process, this is done in accordance with equation (16) below. This will give a more binary final design, but at the risk of introducing some smaller features towards the end. These very small features will however often be of only little significance to the final design and may be removed manually after ended optimization.

2.1.1 Tasks for the designer

Although the benefit of TO is that a design is made without the constraints of the imagination and knowledge of the designer, there are still many parameters that needs to be chosen and which will affect the final structure. Here the most relevant of those will be described to give the reader an understanding of what possibilities the designer has, and by extension what cannot be altered in the existing version of Phazor. As shall become clear, most changes of parameters is a trade-off between several desired outcomes such as functionality, insertion loss, footprint, robustness, fabricability, and calculation time.

Naturally material parameters must be set in the program as well; if nothing else is noted in the text the following has been employed in TO and simulations:

$$\begin{aligned} \epsilon_{air} &= 1 \\ \epsilon_{Si} &= 11.68 \\ \epsilon_{SiO_2} &= 2.085 \end{aligned}$$

Starting point structure and design domain

Obviously deciding on the structure for the optimization and making the specific problem formulation is up to the designer. An important element in this part is size, although compactness is desired in PICs there is a trade-off between making a design smaller and the functionality of the device, this issue is investigated more in depth in chapter 6. This means that although the device is wanted to be small it might be worthwhile to make optimization attempts for a slightly bigger device to begin with. Manually making new iterations with the knowledge from the designs altering the next starting point structure is an option for finding compact, high-performance devices as is touched upon in section 5.2.

Resolution is also a relevant parameter to consider while constructing the starting point structure. A finer resolution tends to lead to a lower insertion loss, however it will increase calculation time and may render fabrication of the final design more complicated while lowering the robustness to fabrication errors. Resolutions that have been tried throughout this project includes 16 nm/pix, 20 nm/pix, 32 nm/pix, 40 nm/pix, 64 nm/pix, and 80 nm/pix. Unfortunately it is currently necessary to employ the same resolution for calculation son the entirety of the structure. A future implementation could be to allow for a finer resolution in the more sensitive parts of the device, so as to keep the calculation time low while maintaining good performance.

The design domain is another part of the input to Phazor which is of great importance. Once more there is the compactness issue. The smaller the DD the shorter the calculation time of the optimization. Running several iterations of a design, it may be beneficial to make a smaller or fractured DD to only include areas where change has previously been observed to have a high impact on the functionality. Naturally this method can also be used to guide the optimization in a specific direction, for example if one local minimum is suspected to exhibit a better performance than another. Although several designs would intuitively exhibit symmetry in the structure, it is currently not possible to enforce it in the TO. To be able to obtain designs that are close to symmetric, it is necessary to ensure a symmetric placement of the DD and objective in relation to the pixel grid. However, it is found that although the emerging designs are near symmetric small deviances occur. This behaviour follows from the issue of converging to a local optimum, and when a design which should be symmetric is not, it is certain that the global optimum has not been found. Enabling symmetry enforcement in the software would be beneficial because of this and because it would lower the calculation time.

Filtering - size and grey scale

Several types of filtering can be included in the TO. They are concerned with the feature sizes and the grey values intermediate between silicon and air in the DD.

The filter regulating the size is a Heaviside filter, as mentioned above. It allows the designer to specify a radius which is the desired minimum value for any feature in the design be it silicon or air. In contrast to when changing the pixel size, this is not a hard limit meaning that some features might end up smaller. A comparison of the results of changing filter and resolution, respectively, is shown in figure 2.1.1. The softer demand posed by the filter is in part due to the change of the filter size throughout the optimization. The usage of this filter along with its benefits and drawbacks are investigated in chapter 6.

The grey scale filter develops throughout the optimization. In the initial phases of the optimization it is necessary to allow for the grey areas for the structure to be able to change and converge upon an overall design. Once the main features of

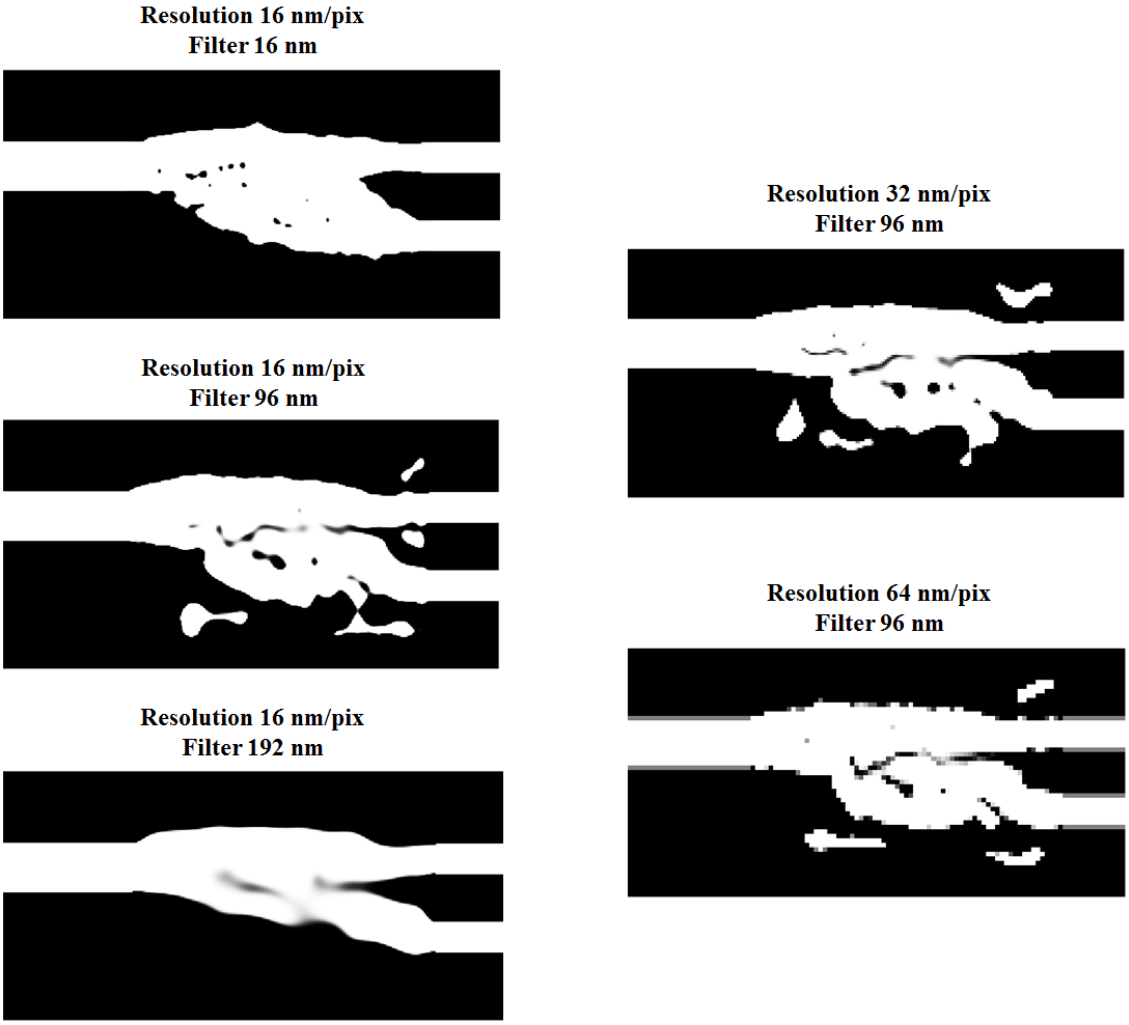


Figure 2.1.1: Examples of the same design, the 2-mode (de)multiplexer discussed in section 5.3.1, presented with different filter sizes (left column) and different resolutions (right column).

the design have been determined the remainder of the optimization process is to refine that structure, this includes removal of the grey areas. Figure 2.1.2 shows an example of the convergence of a design correlating the objective function to the stages of the device. The tolerance of the grey areas is diminished throughout the optimization in accordance with the following equation:

$$\beta = \beta_{min} + (\beta_{max} - \beta_{min}) \times \frac{e^{k \times x} - e^{k \times 0}}{e^{k \times 1} - e^{k \times 0}} \tag{16}$$

$$= \beta_{min} + (\beta_{max} - \beta_{min}) \times \frac{e^{k \times x} - 1}{e^{k \times 1} - 1} \tag{17}$$

where x is the position in the direction of propagation, and β_{min} and β_{max} are the start and end values of β respectively, while k is a constant indicating how quickly the change from one to the other takes place. Generally beginning with a high β_{min} means that only little change is allowed in the beginning, this can for example be very useful if refining an already optimized design, but regularly a low standard-value of $\beta_{min} = 12$ has been employed. A high $\beta_{max} \approx 200$ leads to quicker convergence. The constant has usually been in the vicinity of $k = 2.5$ or at least the same order of magnitude, and minor changes have very little effect on the outcome. Along with this a threshold is set, indicating whether a grey tile should move

towards being air or silicon depending on its grey scale value, this threshold is customarily set to 0.6.

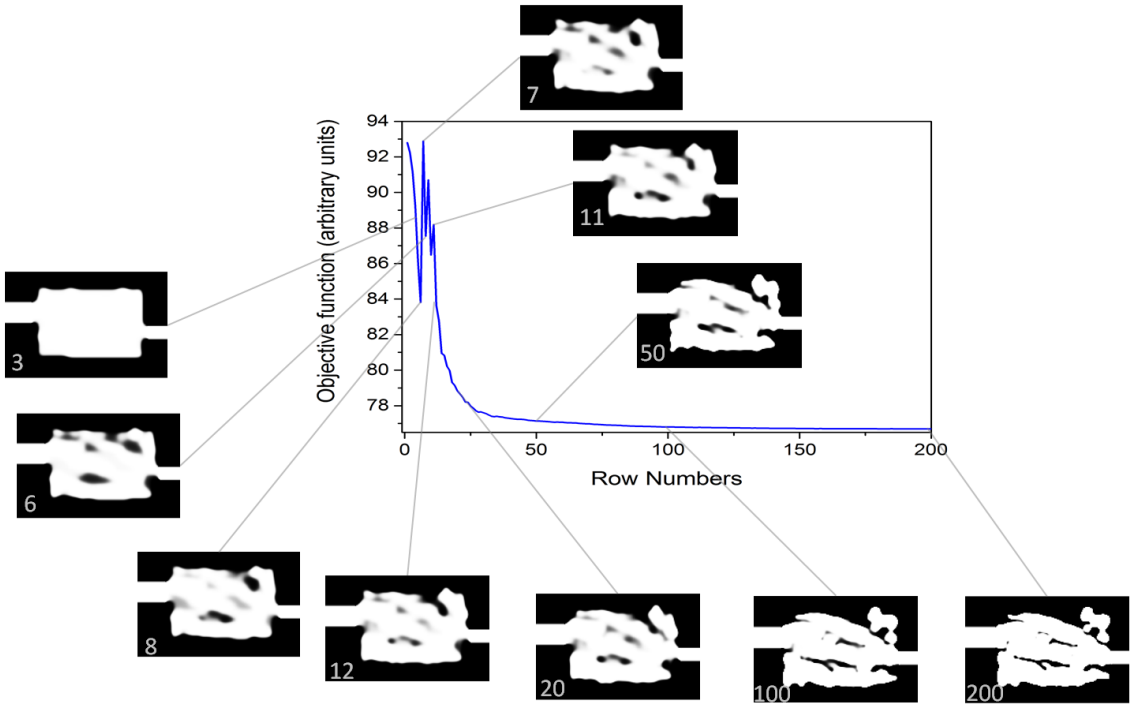


Figure 2.1.2: Illustration of the development of a design and its convergence. The objective function shows the overall convergence. The design is shown at choice points throughout the optimization (iteration number noted on the image) to illustrate the overall changes. In the beginning big changes occur in only a few iterations and large grey-scale areas are present. Towards the end the majority of what happens is refinement and a change to a black/white scale.

Sensitivity and objective weights

Sources and objectives, which are of course paramount to the optimization, is described in the following subsections. However a general thing of note is that every objective is assigned a weight. This is a value which indicates, compared to other objectives, how important this one is to fulfil. In many cases it is beneficial to have all weights equal, as unimportant objectives should be left out. Giving Phazor more freedom will give a simpler design and fewer objectives leads to a shorter optimization time.

The numeric values of the objectives as well as the number of sources and their intensities all influence the sensitivities of the design. The numeric values of the weights should thus be adjusted to reveal sensitivities in the range of approximately 0.01-500. Too low sensitivity values will lead to no change in the design, while too high values will lead to many violent changes in the design as it jumps between local extrema and often the final structure will either not reach convergence or it has too big features to obtain the desired functionality. Examples of how different values of the sensitivity influences otherwise identical designs are shown in figure 2.1.3.

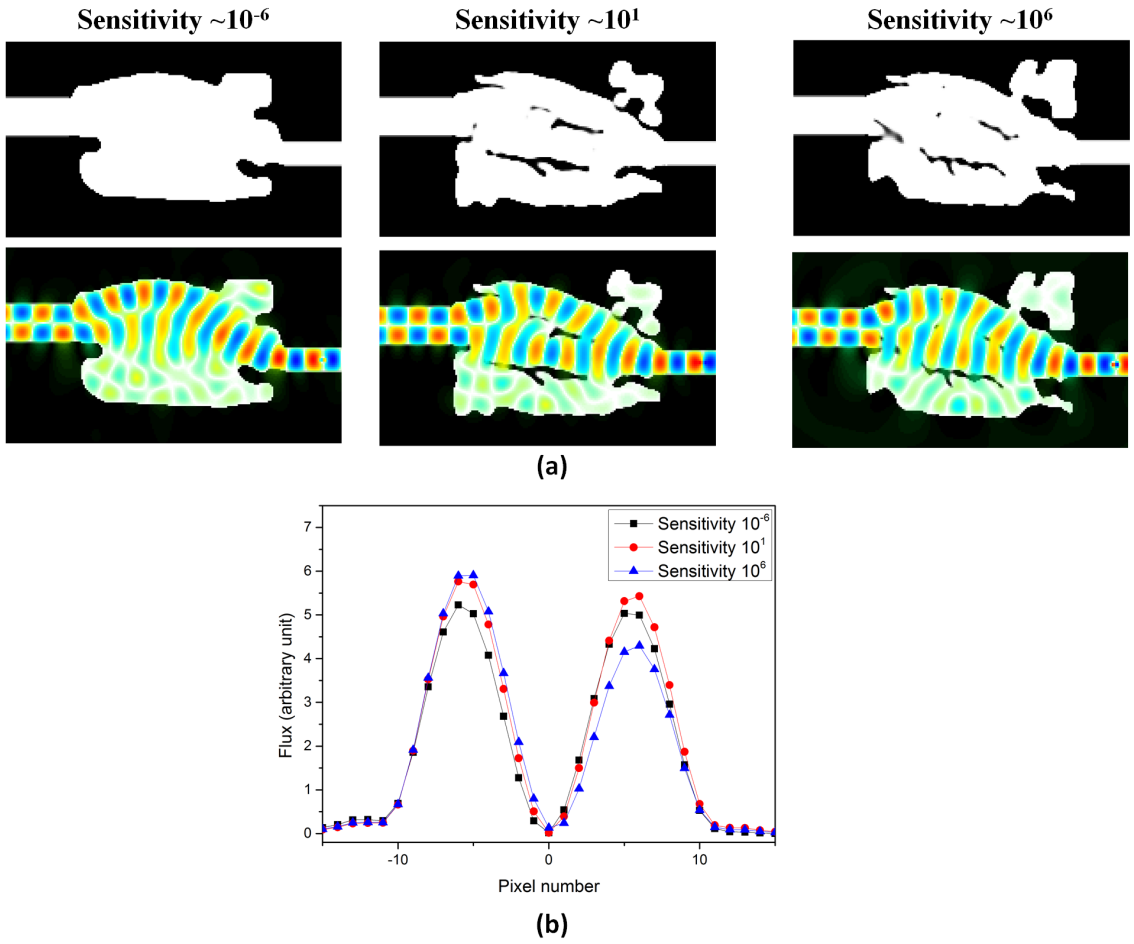


Figure 2.1.3: (a) Examples of designs obtained from the same starting point but with sensitivities in different value ranges due to different values of the objective weights. Upper row shows the bare designs, lower row the design overlaid by the H_z-field to illustrate the impact on the functionality. Design are shown with too low a sensitivity ($\sim 10^{-5}$), appropriate sensitivity levels ($\sim 10^{1}$), and too high a sensitivity ($\sim 10^{8}$). The structure is a mode converter as presented in section 5.2. (b) The mode profiles obtained from the simulated flux after conversion to the TE_1 showing the impact of different sensitivities.

Point energy objective

The point energy objective (PEO) is, from a designer perspective, the simpler of the two objectives. It is placed in a given pixel, and the energy density in that position can be maximized or minimized at that point. Furthermore it is possible to specify the desired center frequency and bandwidth of the resulting pulse. The problem formulations of this type of objective is thus limited to mini- or maximization. It is, however, possible to specify, for example, a splitting of wavelengths, as described in chapter 6. A benefit of this objective is that there is no need for estimating parameter values aside from the weights.

The source which is excited when using this objective is always Gaussian where the amplitude, center frequency, standard deviation (in time domain) as well as cut-off and peak values can be set.

Pulse shape objective

The pulse shape objective (PSO) allows the user to specify what type of pulse is desired centred at a single pixel. This is done by specifying the envelope of the pulse through a function. Most often a Gaussian is stated, in the code, by the form of:

$$env = A \exp\left(\frac{-(t-t_0)\times(t-t_0)}{2\times T\times T}\right) \tag{18}$$

where A is the amplitude of the pulse, t_0 is the arrival time of the center of the pulse, and T is its temporal width. The source is in this case excited similarly with the addition of identifying the center frequency ν:

$$J = A \sin(2\pi\nu t) \exp\left(\frac{-(t-t_0)\times(t-t_0)}{2\times T\times T}\right), \tag{19}$$

The benefits of this objective is the possibility of defining other pulse shapes, and with any shape controlling the time of arrival. This can be useful when dealing with splitting components where various signals can, during optimization, be differentiated by their arrival time. This usage has for example been employed with the mode multiplexers described in section 5.3. A complication does however arise using this objective in regards to the time: It requires an estimate of the arrival time of the pulse. A good first guess can usually be made using the arrival time of the pulse through the unoptimized structure. Meanwhile this will often need to be adjusted for a better guess through several iterations of the optimization. The pulse will need time for scattering if it needs to be altered through the structure e.g. by changing the mode, and too short a time will make the functionality impossible. However, for any pulse it is important to not give the optimization too long a time delay as this will lead to Phazor incorporating unnecessary scattering, and thereby unnecessary loss, in order to delay the arrival time of the pulse to match the objective. Similarly an estimate of the amplitude of the pulse is needed, however this is less of a complication as generally no repercussions arise from overestimating the desired pulse amplitude.

Although this objective has been made to allow for the specification of the pulse shape at a given position it is also possible, for simple structures to use it for maxi- or minimization similarly to what is achieved for the PEO. Using the case of maximization as an example, at the objective point, a pulse with a very low amplitude should be specified and the sign of the objective weight is reversed. In this way Phazor is instructed to attempt to make a pulse completely different from what has been set, and thus not allow for a pulse with a low amplitude. For simple fields such as a fundamental mode, this can lead to a better result, as it is not dependent on properly estimating the value of the amplitude, instead giving Phazor more freedom.

The PSO is a strong objective that allows for versatile design problems, however it does require more fine-tuning from the designer due to more parameters involved in the objective.

2.2 Other optimization methods

There are many methods of optimization that has been employed for the design of PICs. Although TO has been the main focus of this thesis a more Edisonian approach has been used in the work with IBM. Here trial and error, building on a theoretical foundation and previous work, was employed to make design decisions for simulations. These were then used as the basis for choosing what structures to fabricate. The procedure will be briefly elaborated upon in chapter 8. Here a few other methods employed for the design of PICs devices will be cursorily mentioned although it is outside the scope of this thesis to provide a proper review and thus it does not offer a fair comparison to TO.

Optimization methods can be divided into the more classical approach, here called search optimization, and the inverse optimization methods which TO belongs to.

Search optimization allows for a smaller degree of freedom than inverse designs as it will fix specific parameters. This could for example be a restriction that all features have to be round, allowing optimization only of position and radius. In a search optimization these specified parameters will then be altered, within a specified range, to map out the performance of different design configurations. The main difference from the inverse method is that here the focus is on changing the designs systematically and then afterwards evaluating the resulting structures. In contrast the inverse method focuses on the desired result and tries continuously evaluates how any change contributes to obtaining this. Among the benefits of the search method is that by mapping all possible configurations, the designer can be sure to have the global optimum. Furthermore by imposing restrictions on geometries and dimensions, it is possible to ensure designs that will be practical to fabricate. Even though the global optimum is found within these restrictions, it is likely that an even better solution can be found when allowing for more parameters to be changed.

Another substantial challenge that the search algorithms face is the calculation time. When all configurations, including those with a comparatively poor performance, are tested, the necessary number of calculations will be high. Several different algorithms exist to deal with the issue of the computation time a few of which are mentioned here: Wannier basis modelling collects multiple structure points in larger domains whereby the number of points to alter is lessened [17]. The shape optimization method relies on firstly performing a coarse parametric search to ensure that the starting point for a fine-tuning, using a generic direct search algorithm, is close to an optimum [18]. Genetic algorithms are an alternative that does not map out all solutions but rather emulates evolution by intelligently iterating on and combining randomly chosen starting points. At the same time structures which lead to poorer results are discarded minimizing the risk of going in that direction [19].

As mentioned, inverse design methods focus on the desired functionality rather than mapping out different structure variations. This means that all inverse methods tend to define some figure of merit which it is the intent to optimize. Although inverse design is a widespread technique in the field of mechanics, it has only received moderate attention in the field of optics in the last decade. Using an inverse design method, named the objective first method a number of different devices have been suggested [20, 21]. In the last few years interest in inverse design has increased greatly due to a high profile publication using the objective first method [22] simultaneously with one using a nonlinear optimization algorithm [23]. The objective first method is described as being performed in two steps whereof the first looks

for a solution that does not need to fulfil Maxwell's equations. This should then create a refined starting point for a subsequent optimization that tunes it so that it becomes a structure functional within the laws of physics. This second optimization uses the steepest descent method which is also utilizing an adjoint set of equations. In this thesis devices similar to those published in [22] are investigated in chapter 6. Presumably the benefit of the objective first method is that the first optimization step will reveal a structure close to a good local optimum, perhaps even the global one. A known drawback of the current implementation of the objective first method is that it optimizes only for a single wavelength. Broadband functionality is then obtained by merging several devices optimized in several iterations for different wavelengths. Topology optimization does have the ability to optimize for broadband functionality, which in turn may add some robustness to the structures. Neither method does however have any inherent robustness ensuring features.

Other inverse design methods include the direct-binary search [23] and the binary particle swarm [24]. The direct-binary search does, as the name implies, have much in common with the classical searching algorithms. Here iterations are made starting at a randomly chosen pixel and then trying all configurations by altering every single pixel point to be either silicon or air. The change is maintained or discarded depending on the development of the figure of merit. This method has a very high computational cost. Furthermore by working with large square regions as the pixels, one has to expect some deterioration of the functionality when the structure is fabricated caused by the rounding of the structures' edges.

The binary swarm inverse method is similar to the direct binary search in that it imposes the constraint of large square regions and altering the binary values individually [25]. This method does however have a built in search for local optima. Each individual pixel point tend to move towards the value which has, this far, led to the optimum figure of merit that this pixel has 'personally' encountered. A modification according to its neighbours' positions can be implemented. Although both of these methods have divided the structures into square blocks with sizes defined by fabrication limitations it was found to be beneficial to perform an initial optimization with a finer grid and then subsequently modify it to fulfil the fabrication requirements by filtering out small features. It is however suggested to be beneficial to implement filters applied during the optimization, as has been done in TO, rather than only filtering after ended optimization.

No direct comparison has been made of the inverse methods and as of yet no claim of either being better can be staked. Among recent publications TO and objective first have shown better performance but at the cost of smaller features which make fabrication more difficult. In the cases of the direct binary search and the particle swarm it has been a concious choice, which accommodated the calculation heavy methods, to look at lower resolutions to enable fabrication. In all cases variations in the specific designs will occur due to the local minima and exact convergence of the complex problems. The issue of getting trapped in a local optimum, rather than finding the global one, is one of the biggest drawbacks of the inverse design tools. All methods have shown to produce functional devices with small footprints, but lowering the geometrical constraints and the footprint leads to high demands in regards to fabrication. An official comment on the specific Nature Photonics publications using objective first optimization can be found in a letter to the editors in reference [26].

CHAPTER 3

FABRICATION OF PASSIVE NANOPHOTNIC STRUCTURES

Fabricating nanophotonic structures with feature sizes down to just a few tens of nanometres is a demanding process. It has, however, been done for at least a decade in the nanophotonic devices group of DTU Fotonik. It is not the focus of this project to optimize these processes. In this chapter the standard procedures employed at DTU Danchip for the fabrication of the structures presented in chapters 5 and 7 are described. All structures have been made using electron beam lithography (EBL) to be able to obtain the small feature sizes. However from the perspective of economic compatibility, as mentioned in chapter 1, it will be necessary to make designs that can be fabricated with deep-ultraviolet (DUV) [27] lithography (feature size 250-350 nm) or nano-imprint (feature size < 100 nm) which could be compatible with the CMOS fabrication procedures. However other fabrication methods are outside the scope of this thesis.

Every design has been made in SOI material with a 340 nm layer of Si on top of 2 μm SiO_2 on a substrate of Si. While the pattern was defined using EBL, it was transferred from resist to silicon through the use of reactive ion etching (RIE), which will be described at the end of the present chapter.

It should be noted that the structures described in chapter 8 were fabricated at the local facilities of IBM T.J. Watson research center, that I did *not* personally have access to. These processes will not be described in any greater detail.

3.1 Process Overview

The sample is an approximately 2 cm $\times$ 2 cm piece of SOI cleaved from a larger wafer. It will initially be coated with a layer of positive electron beam resist. The design can then be defined through EBL, where the exposed areas are rendered soluble and will be removed during the subsequent development. The pattern is then transferred to the Si through a dry-etch process where the unprotected Si is removed through ion bombardment. At this point the design should be inspected by scanning electron microscopy (SEM) to ensure that the transfer was done as intended before further processing is commenced. A second EBL is subsequently performed to add polymer waveguides on top of the inversely tapered ends of the waveguides leading to the devices. This is done to lower the losses of butt-coupling from fibre to waveguide without an angle to minimized reflections [28]. The polymer waveguides are big enough that the alignment can be verified through optical microscopy before the

sample is cleaved and ready for characterization.

Figure 3.1.1 illustrates the major steps of the fabrication in a schematic view of the sample.

Below the steps of the fabrication of nanophotonic structures in SOI material are listed. Information is somewhat cursory and it should be noted that the values indicated are those used during this project, however they might need to be changed if working under different conditions or with equipment outside DTU Danchip.

Design preparation The design files are prepared with the *L-Edit* software from Tanner Eda and saved in GDSII format.

Wafer cleaving As the designs are not fabricated on wafer scale, smaller chips of about 2 cm × 2 cm will be cleaved before processing. This is done with the wafer coated by a resist (any that is cheap and abundant) which will protect the wafer from dirt and contamination during the cleaving process. The resist is stripped before proceeding and the chip cleaned with acetone, ethanol, and isopropyl alcohol (IPA).

Resist spinning The e-beam resist used was ZEP520A diluted 1:1 in asinole (methoxybenzene) to reduce the viscosity; a layer of ~ 100 nm is sought. This is achieved by spinning 250 μm in two stages: 500 rpm at 200 rpm/s for 5 s. Followed by 6000 rpm at 2000 rpm/s for 60 s.

Pre-baking The resist is pre-baked for 3 minutes at 160 °C, to drive out the solvent and harden the resist.

Electron beam lithography The design files are prepared and the pattern exposed with a current of 0.8 nA.

Resist development The resist is developed for 1:30 min in ZED-N50 and rinsed in IPA for 00:30 s.

Post-bake The resist is hardened through a post-bake for 3 minutes at 110 deg C.

Reactive Ion Etch The pattern is transferred to the silicon through inductively coupled plasma reactive ion etching. Subsequently any remaining resist is removed through sputtering.

SEM inspection Before the waveguides for coupling are added to the structure the sample should be properly inspected, to see if the design was transferred as intended.

Resist spinning An $\sim 4\,\mu$m layer of SU-8 200 resist is spun on top of the etched structure. 250 μL of the resist is spun for 8000 rpm at an acceleration of 4000 rpm/s for 60 s. The chip is subsequently baked for 03:00 min at 90 $<^{\circ} C$.

Electron beam lithography The simpler structure of the polymer waveguides are defined in the SU-8 through EBL with a current of 0.2 nA and without proximity error correction.

Development The SU8 is developed for 2:00 min in PGMEA (MR600).

Sample cleaving The sample is cleaved on the two sides of the design, across the SU-8 waveguides to give clean facets for coupling.

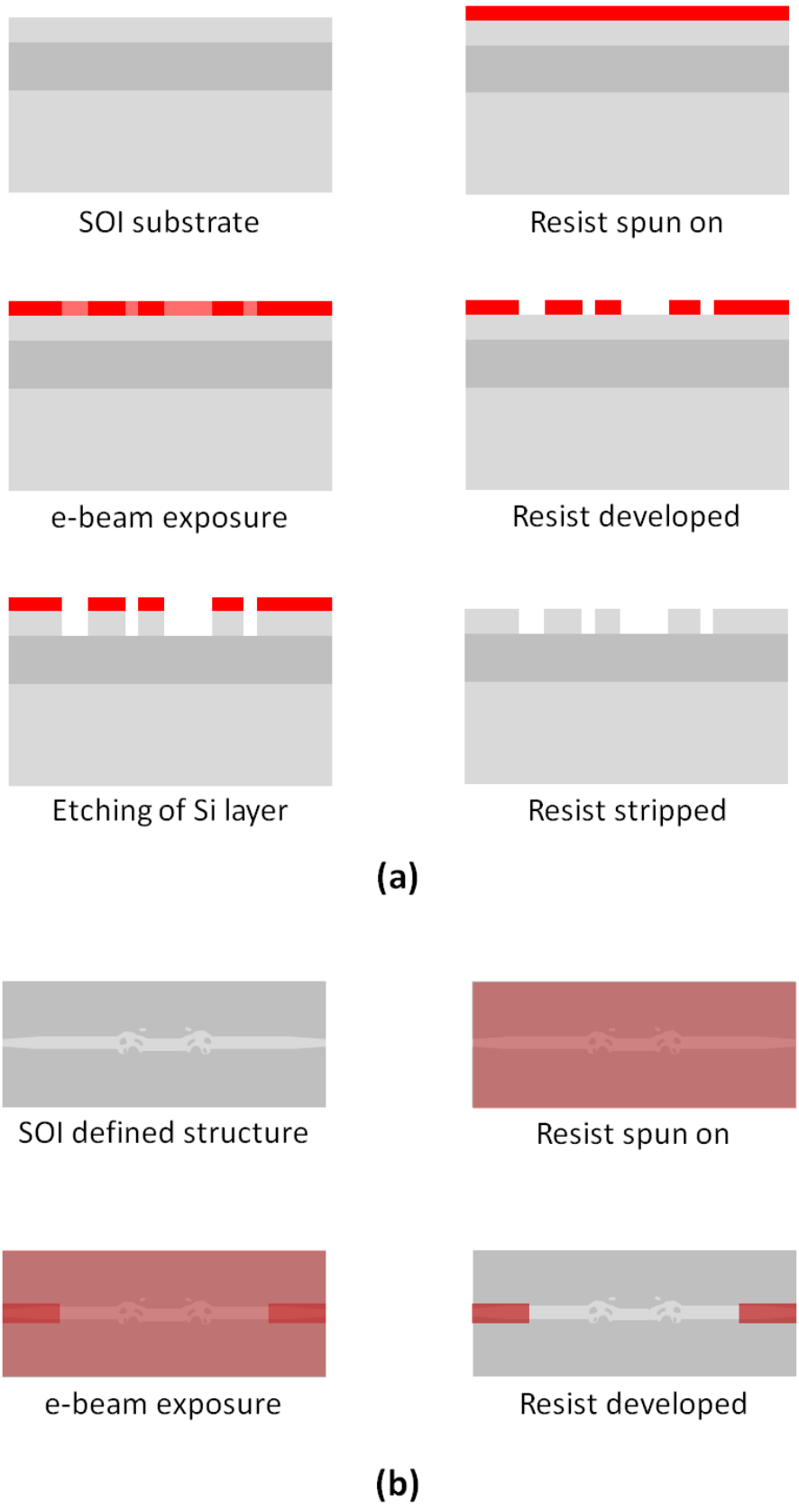
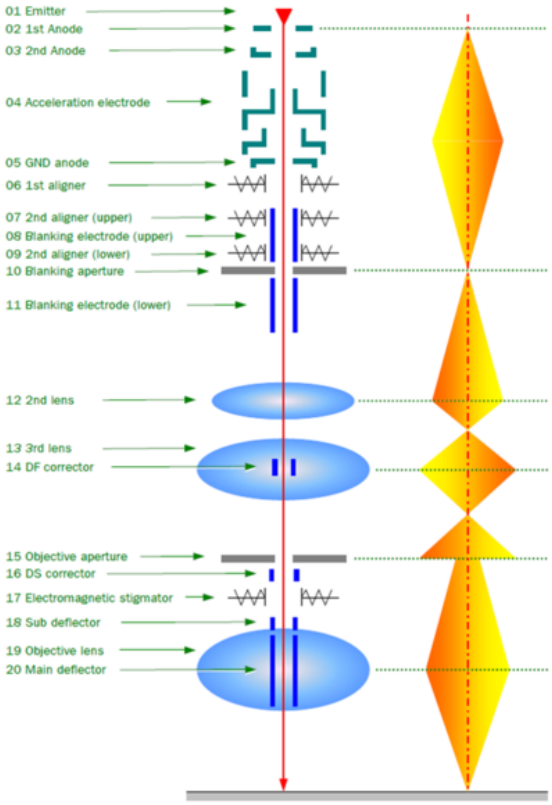


Figure 3.1.1: Schematic overview of the steps of fabrication of nanophotonic structures using electron-beam lithography. (a) Side view of the initial steps of the procedure including the fist e-beam and the etching. (b) Top view of the process of the second e-beam lithography for the application of SU8 waveguides intended for easier coupling.

3.2 Electron beam Lithography

The Danchip Cleanroom electron beam (e-beam) facility is (at the time of writing) a JEOL JBX-9500FS electron beam lithography (EBL) system. It is a spot e-beam system generating the electrons from a ZrO/W emitter. The beam is controlled by a four-stage focusing lens system illustrated in figure 3.2.1. The system can pattern structures with a minimum resolution of 12 nm, it has a scanning speed up to 100 MHz, an acceleration voltage of 100 kV. The maximum field sizes that can be achieved without stitching is 1 mm × 1 mm.



No.	Name	Description
01	Emitter	ZrO/W thermal field emitter
02	1st anode	An electrode that extracts electrons from the electron source.
03	2nd anode	An electrode that extracts electrons from the electron source.
04	Acceleration electrode	For acceleration to 100 kV.
05	GND anode	An electrode at a potential of 0 V. The potential between the acceleration electrode and the GND anode is the accelerating voltage.
06	1st aligner	A coil that aligns the axis of the electron beam from the electron source down to above the blanking electrodes.
07, 09	2nd aligner	Coils that align the axis of the electron beam. The 2nd alignment performs the horizontal alignment and the tilt alignment. The horizontal alignment passes the electron beam through the blanking aperture. The tilt alignment aligns the beam axis below the 2nd lens.
08, 11	Blanking electrode	These electrodes turn the beam on and off. With their 2-stage electrostatic deflection, the blanking electrodes do not shift the beam position on the substrate even when they turn the beam on or off.
10	Blanking aperture	This turns the beam on and off.
12	2nd lens	Together with the 3rd lens, the 2nd lens is called the zoom lens. The 2nd lens determines the beam current. With its zoom-lens feature that sets the 3rd lens value in conjunction with the 2nd lens value, the 2nd lens performs the reduced projection of the 3rd crossover on a fixed point.
13	3rd lens	Together with the 2nd lens, the 3rd lens is called the zoom lens. The 3rd lens determines the beam current. With its zoom-lens feature that sets the 2nd lens value in conjunction with the 3rd lens value, the 3rd lens performs the reduced projection of the 3rd crossover on a fixed point.
14	DF corrector	This corrects for deflection curvature that arises when deflecting the electron beam.
15	Objective aperture	This has six holes with diameters of 60 µm × 3, 100 µm, 200 µm, and 330 µm, enabling you to select one according to the writing conditions.
16	DS corrector	This corrects for deflection astigmatism that arises when deflecting the electron beam.
17	Electromagnetic stigmator	This is for astigmatism correction at the center of the optical axis of the electron beam.
18	Sub deflector	An electrode that scans the electron beam.
19	Objective lens	A lens that projects the 3rd crossover on the substrate to form a spot beam.
20	Main deflector	An electrode that positions the electron beam.

Figure 3.2.1: schematic of the electron beam lithography system's tower, indicating the active elements and their usage. Directly from [29].

The e-beam system allows for patterning of a user-defined structure in either a positive or negative photoresist. There are few limitations to the geometries that can be defined, aside from the resolution there is also a limit to the number of polygons that the software can handle. It is essential that the correct operation parameters are chosen to obtain proper pattern transfer and shortest possible writing times. The first and foremost parameter is the clearance dose also known as the resist sensitivity, Q. This is the amount of electric charge which is necessary to deposit into an area on the sample in order for the resist to develop correctly. The clearance dose is a function of the dwell time t, the beam current I and the pattern writing area A as follows:

$$Q = \frac{tI}{A}. \tag{20}$$

The necessary clearance does of the ZEP positive photoresist used in this project is around 200 μC/cm^2. The EBL system has a number of different apertures that can be used, but the beam size is a function of both aperture size and the beam current. The chosen beam current is based on the design feature sizes; increasing the current increases the beam size. For good dose uniformity the step size should

be chosen to be around the size of the beam spot, possibly slightly smaller. The main limiting factor of the dwell time is the scanning speed of the system. However it should be noted that if proximity error correction is applied there is a further limitation. In this case there will be a need for rapid changes in the dwell time which is very demanding of the electronics of the area deflectors of the system. In this case the dwell time should increase, to ensure that extremely short times will not be necessary for any part of the design; thus a lower operating current, typically in the order of 0.2 nA, is employed [29].

Proximity error correction is a technique utilized to account for the changing density of the designs. When the e-beam impinges on the resist or the substrate it will be scattered. The scattered electrons will deposit energy into the resist outside the intended area of the beam. This happens in two different mechanisms the first is denoted as forward-scattering, when the electrons scatter further into the resist leading to beam broadening. The second scattering mechanism is backward-scattering when the electrons bounce in all other directions within the resist or substrate. The effects depend on both the acceleration voltage and the density of the structures. In the systems investigated in this project backscattering is the dominant effect and has a radial range on the order of 30μm from the injection point. This is accounted for in the EBL process by dividing the design area into various sections and differing the e-beam dosage depending on the density of features in that region. In this way areas where many features needs writing is subjected to a lower dose than where there are less structures[30].

3.2.1 Sample design

A number of design parameters were kept similar for all of the various devices to make fabrication easier. The EBL system has a chip cassette that allows for the mounting of chips in the size of a few cm^2 and with beaming slots of 4 mm, 8 mm, 12 mm and 20 mm. For the designs presented in this thesis the two middle size slots were used. As one of the goals of the usage of TO is to obtain compact devices, it is necessary to rout the light to and from these structures. This is done through butt-coupled ridge waveguides that are kept relatively standardized across the chips. To make the butt-coupling easier a leading inverse taper with an SU-8 polymer on top [28] is incorporated. The SU-8 section is 1.7 mm long and is followed by 1-2 mm of ridge waveguide, leading to a full sample length of approximately 7 mm. As the used photoresist, ZEP520A, is positive the waveguides are defined by two trenches with a spacing equal to the width of the waveguide. The waveguide width is tapered from 90 nm in the coupling region to the necessary width of the devices being 500 nm, the trenches themselves have a width of a few μm. Proximity effects of the e-beam procedure should be avoidable at distances over 35 nm, but as a rule of thumb all sensitive and important structures are spaced with at least 50 nm.

In all of the designs, features which are larger than $\sim$ 50 nm are shrunk by $\sim$ 15 nm before the e-beam scribing. This is done in anticipation of line-broadening during etching. This does of course lead to some uncertainty in regards to which feature are shrunk and which are not. Shrinking the smaller features is not possible as they would become too small to actually appear during scribing, however allowing them to be broadened may lead to increase of losses. Currently these choices are not systematic but relies on the judgement of the designer. This is also one of the reasons that design robustness is very important as will be touched upon in section 5.5.

3.3 Reactive Ion Etching

The presently available etching facilities at Danchip used in this project is an Advanced Silicon Etcher (ASE) STS MESC Multiplex ICP no. 30343 [31]. It is not the most state of the art equipment for etching held at the cleanroom, but it is well suited for the needed process as it has been optimized for deep etching of Si. The chamber is supplied with seven different gasses: SF_6, O_2, C_4F_8, CO_2, Ar, N_2 and He, where the latter two are for purging of the chamber and cooling of the backside of the wafer, respectively.

The etching is based on the Bosch process where the sample is cyclically subjected to etch with SF_6 and O_2 alternated with C_4F_8 for passivation. The etch is thus primarily due to free fluorine present due to plasma dissociation of SF_6 [30].

$$e + \mathrm{SF}_x\,(g) \rightarrow \mathrm{SF}_{(x-1)}\,(g) + F\,(g) + e \tag{21}$$

where $x = 6$ in the beginning but the process can continue down to SF_3. The actual etching process then follows

$$\mathrm{Si}\,(s) + 4\mathrm{F}\,(g) \rightarrow \mathrm{SiF}_4\,(g). \tag{22}$$

The electrons allowing for the dissociation of the SF_x is inductively coupled by a coil surrounding the plasma chamber. The high plasma densities of the inductively coupled plasma leads to a higher etching rate. By keeping the camber pressure low the amount of plasma scattering is kept small which allows for a better directionality of the ion bombardment that clears the surface of the substrate of reaction products. The benefit of the cyclical Bosch process is that the protective film is formed separately in intermittent cycles with the etching. The film is removed on horizontal surfaces by the ion bombardment, while it is only very slowly removed from the vertical surfaces due to scattering and during the etch thus effectively increasing the anisotropy of the system.

As the etch rates are very dependent on the chamber environment, and this can change depending on the contaminants from previous etches and the chosen carrier wafer, the chamber is always cleaned with O_2 on a dummy wafer before the sample is inserted. Aside from stripping loose particle from the chamber, this also starts the heating of the chamber necessary for thermal stability of the etching process. Furthermore two test samples will be etched with a different number of cycles to determine the present etch rate. The Si layer present on the test sample is measured before and after etching using a FilmTek 4000$^{\mathrm{TM}}$ from Scientific Computing International, which can handle film thickness in a range of 0.05 μm to 250 μm. After the etch rate has been determined the actual sample can be etched ensuring that all Si removed. The SF_6-etching is very selective between silicon and silica so the underlying SiO_2 layer can effectively be used as an etch stop, hence the cycle number will be chosen to include slight over-etching rather than under-etching. When the etching is done any remaining resist will be stripped by an O_2 cleaning with the sample in the chamber.

Using the ASE leads to a good structure representation and steep side walls, however the etching process will cause some amount of line-broadening which is likely to be increased with over-etching. The diameter line-broadening is less than 40 nm and is accounted for by shrinking the larger features of the designs before patterning it onto the chip.

CHAPTER 4

PASSIVE CHARACTERIZATION METHODS

The devices presented in this thesis have experimentally been subjected to passive characterization only. Their functionality is established based on two parameters their mode profiles, revealing the stability and conversion abilities of the devices, and the transmission spectra illuminating insertion losses and cross-talk. The same base set-up is used for both types of measurement and it is sketched in figure 4.0.1.

Light is coupled from a tunable light source (TLS) by single mode fibre through a polarization controller and to a tapered fibre. Two sources (Ando AQ4321A Tunable laser source) are available with a range of either 1480 nm - 1580 nm or 1520 nm - 1620 nm. The light is butt-coupled into the chip, which is mounted on an x-, y-, z-stage which can be be controlled both manually and by piezo stages. Above the mounted chip an optical microscope, with the possible magnifications of 50x/0.70, 20x/0.40, and 5x/0.13, is placed allowing for the rough alignment. Further an InGaAs infrared camera (IR-Cam – Xenics XEVA XC130) is placed on the top output of the microscope and is used for the mode profile imagining as will be described below. The light which has traversed the chip is butt-coupled out into another tapered single mode fibre and directed to an Ando AQ6317B optical spectrum analyser (OSA). Both the TLS and the OSA can be controlled by in-house written LabView software.

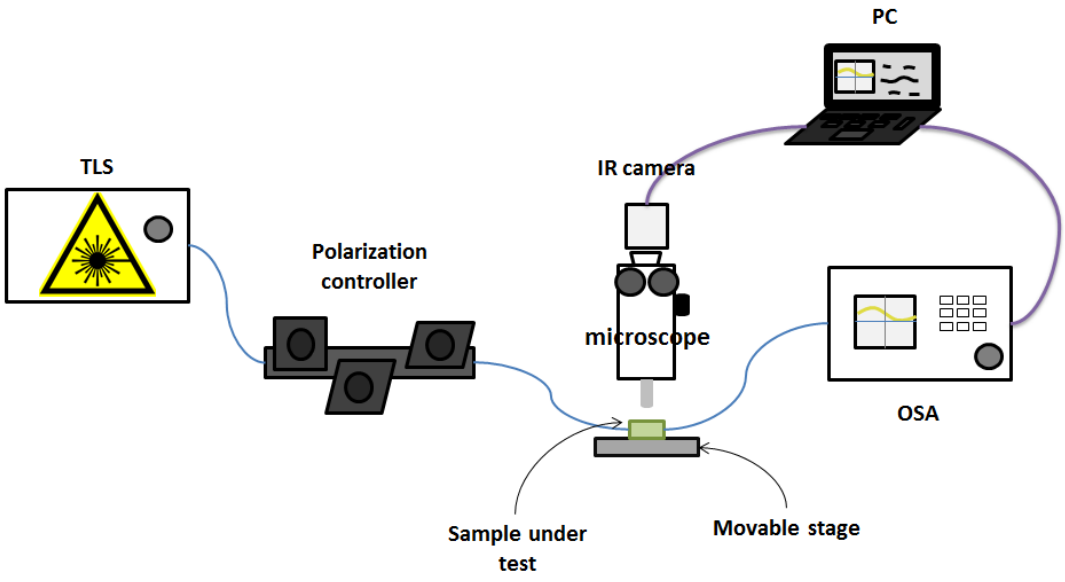


Figure 4.0.1: Schematic representation of the measurement set-up for passive characterization of the silicon integrated devices.

4.1 Mode profile recording

Although not important for all of the structures presented in this thesis, the recording of mode profiles was an important element of the characterization for a number of them. Mode profiles are measured from devices of a configuration as illustrated in figure 4.1.1. The light is butt-coupled to the chip and traverses the device under investigation, the waveguide is then tapered out to a vertical grating coupler [32]. The light, which is coupled out of plane, is collected by the optical microscope and imaged by the IR camera placed on top.

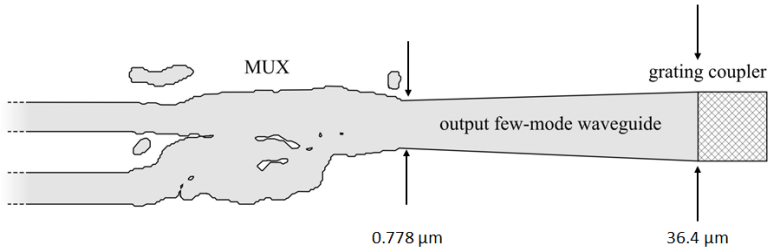


Figure 4.1.1: Illustration of the type of chip design used for measurements of the mode profiles. Here the sketch depicts a mode multiplexer with two single mode inputs and a multimode output expanded to the grating coupler. Structures not to scale.

The mode profiles are a great method for qualitatively illustrating the functionality of a device. Meanwhile, the set-up does have a number of challenges for more quantitative analysis. The extinction ratio of a mode can be determined from line scans across the recorded mode profiles. There are however some challenges in regards to obtaining the correct value for the extinction ration, namely saturation and image resolution. If too much power is injected into the taper, the peaks of the modes will be saturated, this causes a flat top in the line-scan and leads to a lower extinction ratio than actually exhibited. The resolution of the camera may lower the valley of a TE_1 mode for example, as the high gradient region and the valley point are resolved by too few pixels. In this manner especially the bottom point may be the value of an average of the area close to the valley rather than the actual minimum which will lessen the estimate. This resolution limitation can be helped by increasing the magnification to its maximum. Figure 4.1.2 illustrates these different cases of limited an best-case mode profiles recorded by the set-up.

The mode extinction ratio is calculated from the integrated powers of the present modes. However the peak extinction ratio (PER) can be found simply as the difference between the level of the lowest peak of the mode and that of the most shallow valley. Although not as precise it is a good estimate and it can be determined without much post-processing. Calculating for the best case mode profile presented in figure 4.1.2 (black) it is found that $MER = 20.90\,\mathrm{dB}$ and $PER = 16.80\,\mathrm{dB}$ underlining the closeness of the results and that the PER will be a conservative estimate.

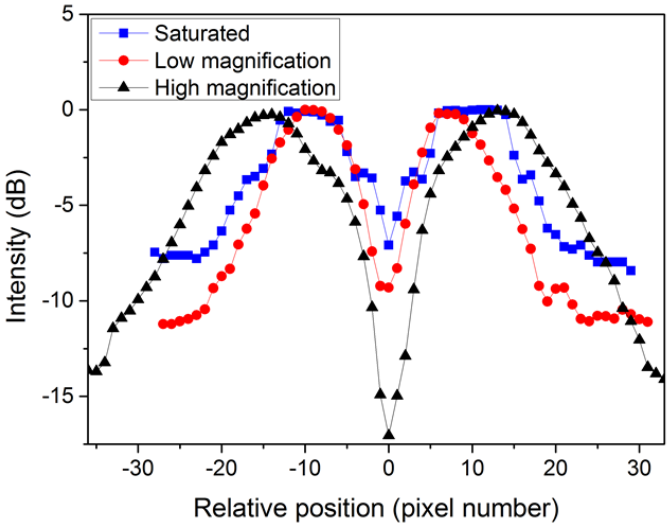


Figure 4.1.2: Example mode profile line-scans used for determining the extinction ratio of a TE_1 mode. Three examples illustrate the issues of saturated signals (blue) and low image resolution (red) along with how a higher magnification significantly increases the measured value (black). Due to the change in resolution the width and position of the last dataset is shifted, it does however clearly illustrate the difference in the extinction ratio. Profiles are from the 2-mode multiplexer device presented later in section 5.3.

4.2 Insertion loss measurements

in the current set-up only fundamental mode light can efficiently be coupled to the waveguides. Because of this mirrored structures, as illustrated in figure 4.2.1, are used for measurements of the insertion loss. As the investigated designs are reversible and characterized in the linear regime, the insertion loss of a single device can be found by halving the values of the recorded spectrum. A design challenge is posed in determining how long the light should propagate before the reverse design is presented. Generally the measured loss increases with this length, so letting the light propagate for longer distances will lead to the measured loss not accurately reflecting the loss of the device under investigation. However the design cannot be considered functional if the light does not efficiently propagate after traversing it. Different lengths between reflected structures have been investigated in this project and, although giving the opportunity for interesting results, it does complicate comparison across designs. No matter the length, investigating the insertion losses using these mirrored structures will lead to conservative loss estimations for a single device.

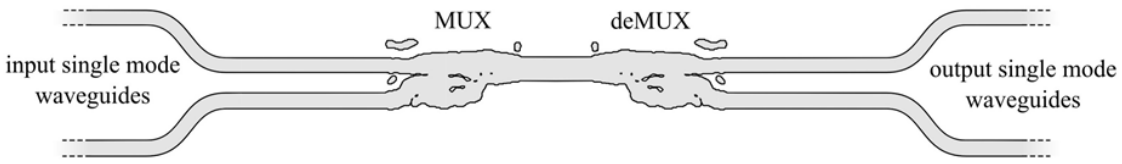


Figure 4.2.1: Illustration of the type of chip design used for measurements of the insertion loss. The structures are mirrored due to the necessity of the light coupled on and off the chip being of the TE_0 mode.

When coupling light in and out of these waveguides for insertion loss measurements, it is done by manual alignment. First step is done using the optical microscope and the x-, y-, z-stage on which the sample is mounted. Once the rough

alignment has been done the TLS, set on an a wavelength where a high output is expected for the particular design, can be turned on. If a signal is detected, a better alignment is obtained first from the screws of the stage and subsequently by using the piezo stages to get the highest possible signal. Once the fibres are aligned a scan of the full bandwidth of the TLS is made to record the wavelength spectrum.

A complication for this manual alignment is that there is no guarantee that the alignment is ideal and misalignment will influence comparison of different designs. A mitigating factor is that the measurements will always be a lowered estimate, thus never claiming a better performance than the device does in fact have. This is of course only true if the reference measurements can be trusted, but more on this below. An additional way to safeguard against alignment discrepancies is by having multiple chips or repetitions of the same device across the same chip, to ensure that similar levels are obtained for all measurements.

All chips are supplied with a number of reference waveguides used for normalization. Various incarnations may be included depending on the devices under test (including bends, tapers, various lengths, etc.) but every chip will as a bare minimum have straight reference waveguides running across the width of the chip as well as a reference including a photonic crystal (PhC) waveguide which will allow for the polarisation to be determined.

There are many risks of variation across a chip including the thickness of the silicon layer, angled cleaving, dirt, damage and uneven etching during fabrication. To safeguard against this every chip has multiple sets of the relevant reference waveguides spaced out along the vertical direction of the chip and placed between the active devices. Figure 4.2.3 shows an example of a chip overview illustrating the reference placement along with the various references measured along the chip. It is evident that there are similarities of the reference level across the various measurements such as a systematic error occurring around 1540 nm and that the signal strength generally drops for higher wavelengths. Even though the reference spectra look similar, there is an overall change in power level for the different chip locations. To give the most conservative estimate of the performance of the devices, the highest measured reference level is generally chosen for normalization. Exceptions to this occur only if deviations are systematic (e.g. power level changes continuously in the vertical direction on the chip) or if a specific spike or drop is seen only in certain regions of the chip - in this case the reference waveguide closest to the measured device will be used.

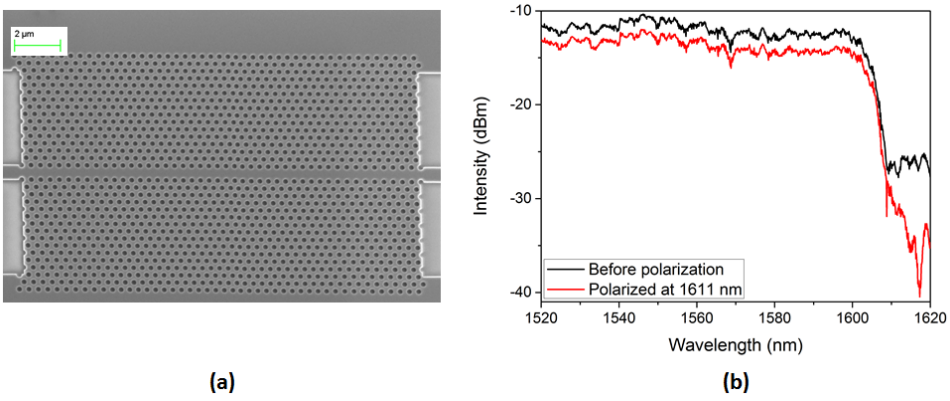


Figure 4.2.2: (a) SEM image of a photonic crystal used as a reference for TE-polarization. (b) Transmission spectra of a waveguide including a photonic crystal section before (black) and after (red) the TE-polarization has been achieved.

Photonic crystals are structures with a periodic variation of the refractive index which give rise to a photonic bandgap [30]. This means that light of a certain wavelength cannot exist within the structure. As the fabricated SOI PhC lack translational symmetry in the vertical direction mode the present modes will not be purely TE or TM. The modes in the PhC will however been TE- and TM-like. As practically only the electrical field propagates along the PhC the strongest cut-off at the photonic band occurs when the mode is as TE-like as possible. This allows for tuning the system to a TE polarization. To do so the fibres are aligned as usual at a low wavelength where the transmission level should be reasonable. A line-scan is performed and polarization optimization done at the wavelength of the cut-off, this time ensuring the lowest transmission power and thus the cleanest cut-off. Figure 4.2.2 shows an SEM image of one of the reference PhCs along with the unnormalized spectra before and after polarization optimization. The TE-polarization should be verified throughout the day when measuring as it may drift over time.

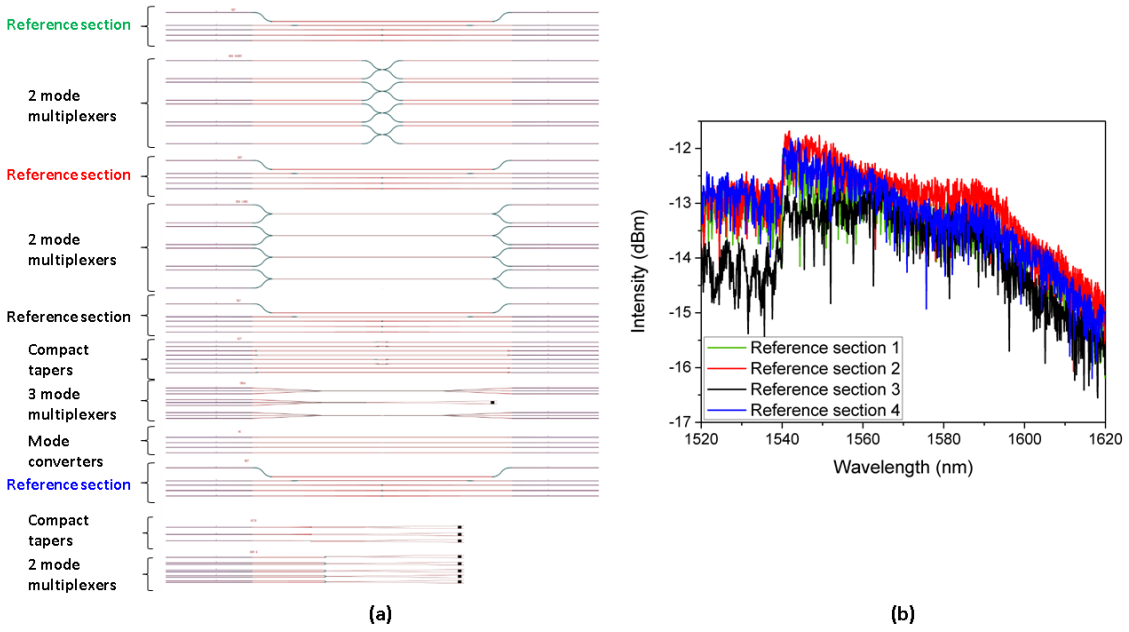


Figure 4.2.3: (a) Example overview of a chip design including four reference sections spaced between the active devices. (b) Transmission spectra of four identically designed straight reference waveguides placed at different vertical positions on the chip.

4.3 Grating measurements

Measurements of the cladding modulated gratings designed and fabricated at IBM and described in chapter 8 were conducted in part at the facilities at the T. J. Watson Research Center and in part at DTU. The set-up for these measurements was very similar to the one described above, but a few differences will be noted here.

As the wavelength range relevant for the designed grating is spaced around 1310 nm, somewhat outside the conventional telecom range, another TLS was needed as light source; here the New Focus swpt wavlength laser TLB 6500-H O was used. Unfortunately this TLS does not have the capability to interface with the OSA and instead a LabView script was made to allow for recording the spectrum with a HP 81536A power sensor detecting in the range of 800 nm to 1700 nm [1]. A challenge with this

[1] Many thanks to Dagmawi Alemayehu Bekele who supplied me with the basis of this script.

set-up is that the power meter needs time for the signal to stabilize leading to overall slow measurements even for a relatively low resolution. The long measurement times for a single scan does pose a risk of system drifting during the course of a single scan. This is especially challenging to evaluate as the loss tended to increase for higher wavelengths regardless. A solution to this became including a few points of measurement for a lower wavelength at the end of the scan, to determine if drift took place.

The measurements of the grating from IBM are made for the sake of determining the reflection strength of the different designs and obtaining low insertion losses were not a concern. For this reason the various devices are not normalized and references are not included in the same way as described above.

CHAPTER 5

MODE DIVISION MULTIPLEXING COMPONENTS

5.1 Introduction

As touched upon in the introduction of this thesis, one of the reasons to investigate the development of optical communication is an increasing demand for data transfer. In the field of fibre optics several technological leaps have been introduced to increase the capacity [3]. It has been suggested that SDM could be the next technology to push the limit further. The high bandwidth for optical interconnects is essential, making the investigation of on-chip SDM is a natural choice as a means for increasing this. Furthermore it is relevant to follow the development of the field of fibre optics to allow for easier interfacing and co-processing. Space division multiplexing can be realized in several different ways. Examples include adding more channels which can be realized for fibres by bundling them or making multi core fibres. Alternatively one can add more modes in a single waveguide. It is the latter method that is relevant to the present project.

The technique of MDM is interesting for PICs devices, regardless of fibres, for the potential of routing multiplexed optical signals on chip, as this can potentially reduce the necessary footprint for the same amount of data transferred. Using MDM rather than WDM also offers the benefit of needing fewer integrated light sources in all-integrated devices. Building blocks necessary to enable MDM include mode converters and mode multiplexers. Other groups are also investigating such devices employing several different schemes. Although it is outside the scope of this thesis to supply an actual review of these various techniques, some of the more frequently employed methods will be outlined in the following sections as relevant to allow for a comparison to TO.

This chapter deals with the design of four different mode-converting devices using TO. The best designs have been fabricated and the experimental data characterizing the devices will also be presented in this chapter. Firstly the simplest device type, the mode converter is introduced. This includes the most base form operating on the TE_0-to-TE_1 modes and a more complex mode converter operating on the TE_0 and TE_2 modes. Following that study the more complicated 2-mode (de-)multiplexer (2MUX) is presented, it too operates on the TE_0 and TE_1 modes. A cursory investigation of the robustness of these particular designs is included here. The final device of this chapter is a demonstration of how the multiplexing can be expanded to include more modes, in this case adding the TE_2 mode to create a 3-mode (de-)multiplexer (3MUX) design.

5.2 Mode Conversion

The first design taken under consideration was the simplest mode manipulation: the conversion of a TE_0 mode to a TE_1 mode. Although it has the simplest functionality, several iterations were made before a suitable design was obtained. As it was the first device created with TO during this project, it did also include much learning about the process and its various strengths and limits.

Numerous groups have investigated methods of making mode converters for PICs. Typical device designs rely on either the phase matching of different waveguides [33] or the scattering effects in a single waveguide [34, 35] .
In the case of optical interference the concept can be described, in a simplified manner, as splitting up the TE_0 signal propagating through a waveguide and then extending one arm enough to delay half the signal to introduce a phase shift of exactly π. When the waveguides are subsequently collected, the mode will become TE_1 [33]. This concept is sketched in figure 5.2.1(a). The designs can be made relatively compact with lengths in the order of $10\,\mu$m, however scaling them up to include other, higher order, modes will greatly increasing the footprint.

The more structurally complex devices based on scattering have the benefit of very small footprints of no more than a few square vacuum wavelengths but it comes at the cost of a narrow bandwidth [34]. The petite features do also make these devices difficult to fabricate and as of yet they have not been experimentally verified. They tend to rely on engineering the photonic bandgap of PhCs, either in the form of holes or rods, by breaking of the structural periodicity and/or symmetry. An example of such a structure is shown in figure 5.2.1(b), however the concept cannot be generalized as other published devices have very different distributions of geometries.

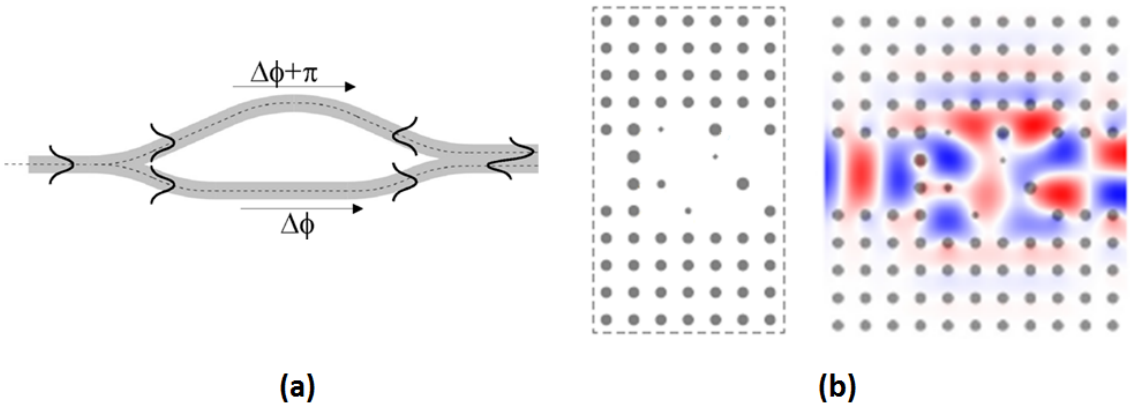


Figure 5.2.1: Sketches of working principles of TE_0-to-TE_1 mode converters as they have been realized by other methods than TO. (a) Using interference between two waveguides and introducing a π phase shift in one of them. Figure directly from [33]. (b) Example of a photonic crystal rod structure which allows for mode conversion on a very compact scale. Figure modified from [35].

In this section various iterations of the mode converter design will be presented. The first step was a design in a PhC and from there we move onto creating topology optimized structures directly in a PhW. Several iterations on the starting point structure is introduced to obtain a functional device. Once a device is obtained further manual changes of the starting point is applied building on the knowledge of the previous optimizations. Finally it is demonstrated, that once the basics of a design is known, it is possible to add more complex functionality.

5.2.1 Photonic crystal mode converter

Topology optimization has been used extensively for the design of PICs at the DTU department already. In the past, the method has however, generally been applied to PhC devices [36, 37]. This was a natural starting point as the PhC already consists of a number of small scattering elements and as it was already well known that various functionalities could be obtained by perturbation of the periodic structures [38]. Hence the starting point of this project was also a topology optimized structure in a PhC, namely a mode converter working on the TE_0 and TE_1 modes. The design was already finished when this project was commenced, and I have thus only been involved in the fabrication and characterization processes of this particular device.

For the starting point of the design a dispersion engineered PhC waveguide was used. By increasing the diameter of the second row of holes compared to all others in the structure, the TE_1 mode's dispersion is made near monotonic and near linear. Topology optimization is applied to a central area in the waveguide as well as to the coupling from the PhC to the PhW, see figure 5.2.2. The optimization was performed in 2D and 3D to evaluate the necessity of the more computationally expensive 3D optimization for correspondence between experiment and simulation.

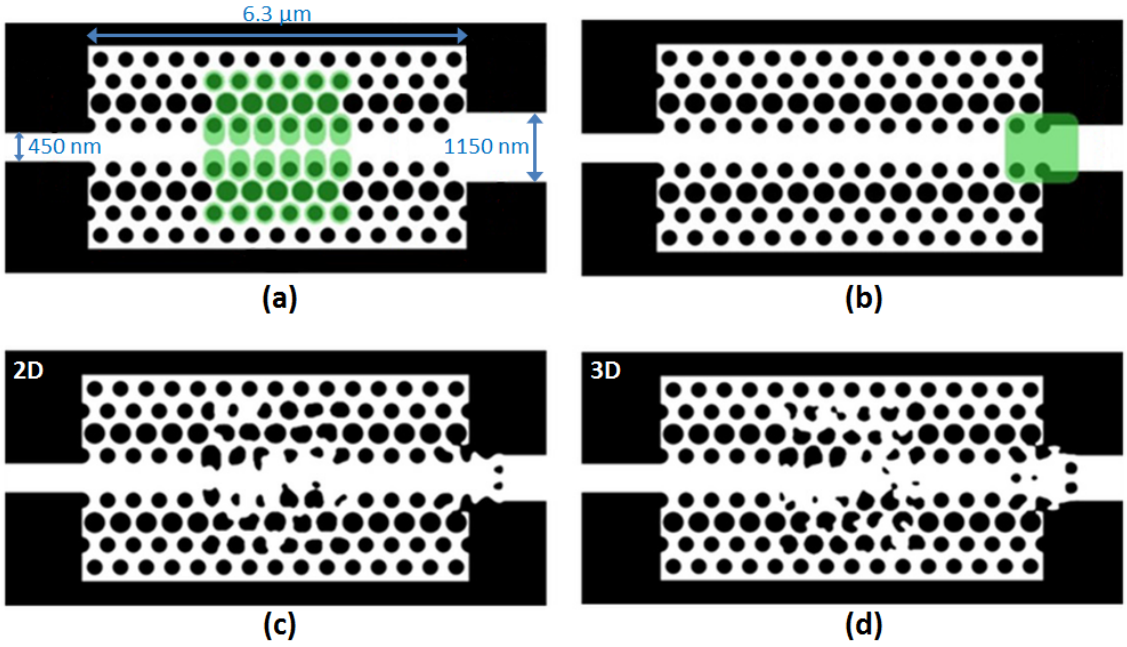


Figure 5.2.2: Starting point structures of optimization of (a) the mode converter design and (b) the coupling from the photonic crystal to the photonic wire waveguide. Designs obtained through topology optimization in (c) 2D and (d) 3D. Figures adapted from [39].

Figure 5.2.3 shows the simulated as well as experimental performance of the topology optimized PhC mode converter. The simulations include 3D FDTD simulations of the devices optimized in 2D and 3D and in the latter case with or without the optimization of the coupling to the PhW. It is apparent that the 2D optimized design degrades when subjected to 3D simulation. This indicates that the full 3D simulation is needed for the best results when realizing the structures through fabrication. Furthermore it is evident that the losses of coupling from the PhC to the PhW can be significantly lessened by optimizing the interface as well. This is essential for the potential of implementing PhC structures in PICs as it would not be feasible to incorporate PhCs on the entirety of the chip. Figure 5.2.3(b) reveals good correspondence between the calculated 3D FDTD spectrum and the experimental results

both quantitatively and qualitatively. The existing discrepancies at the ends of the spectrum are mainly caused by the finite grid size of the simulations. Furthermore the measured mode profiles shown in figure 5.2.3(c) verifies that the conversion does indeed take place, as long as the wavelength is below the photonic bandgap. The top row of recorded mode profiles show the conversion brought on by a single device. The bottom row confirms that mirroring the design, as in the structure of figure 4.2.1, causes the light to be converted from TE_0 to TE_1 and subsequently back to the TE_0 mode which is then recorded.

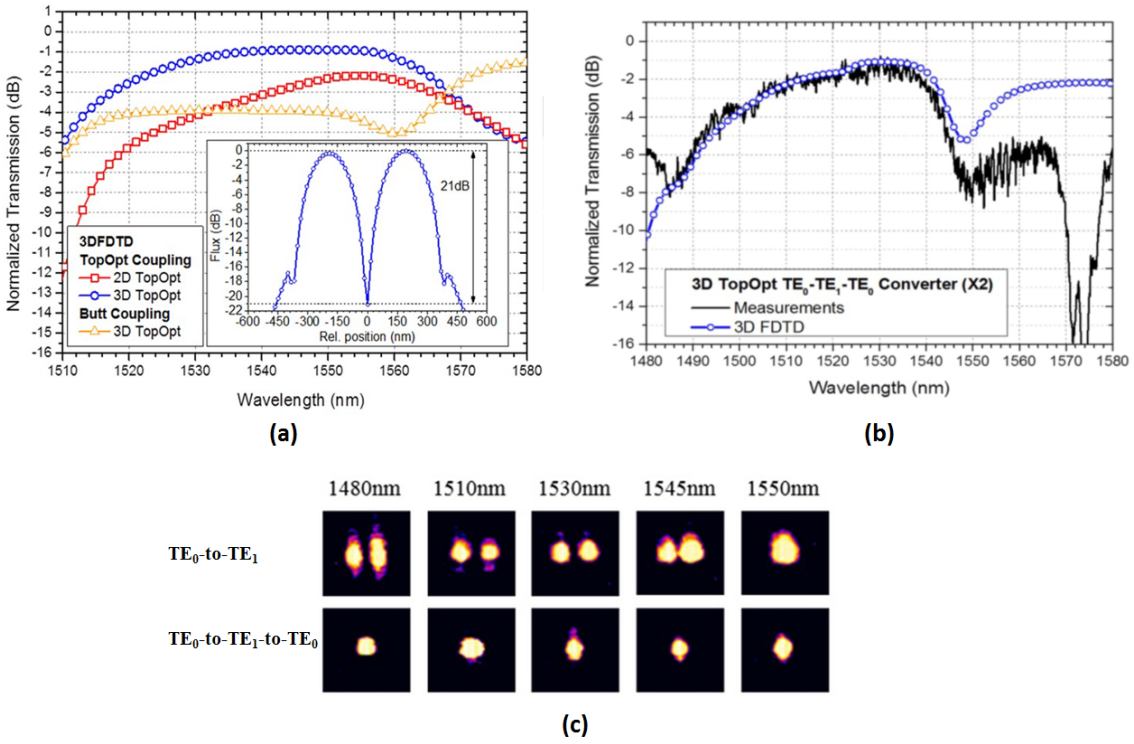


Figure 5.2.3: (a) Simulated spectra of the various optimized mode converter structures in 2D and 3D the latter with and without optimization of the interface between the photonic crystal and the photonic wire waveguide. (b) Measured spectrum plotted in comparison with the 3D FDTD simulations. (c) Recorded mode profiles verifying the conversion of the mode. Figures modified from [39].

These results show that it is possible to use TO for the design of a functional TE_0-to-TE_1 mode converter in a PhC and they lay the foundation for the rest of the work in this chapter. From here on the PhC will no longer be used as starting points for the devices. Rather it is the aim to confirm that it is possible to use an arbitrarily chosen, simple starting point for the TO and still obtain complex functionalities sought after. Ideally, not including a PhC will allow for structures which are more easily fabricated and which require less chip area. Furthermore omitting the PhC will mitigate the need for additional optimization of the coupling between PhC and PhW as the coupling between the device region and the waveguide is included in the same overall design. Having the knowledge of the functional PhC mode converter is however valuable as it shows the feasibility of the design and gives a standard for comparison.

5.2.2 Straight Mode Converter

The goal of the design process undertaken in this and the following sections is to reversibly convert a TE_0 mode to a TE_1 mode. This is the same functionality as obtained for the PhC presented in the previous section, but now the process is sought in a slab waveguide. This was done in a straight PhW tapered from a narrow waveguide with a width of 432 nm, supporting only the fundamental mode, to a wider waveguide of 688 nm supporting also the TE_1 mode. The starting point structure is illustrated in figure 5.2.4, with a design domain of being 3.12 μm $\times$ 1.68 μm (yellow).

Three configurations of the PEOs were attempted for the optimization as marked in figure 5.2.4(a). (B1, lightest purple, square) was a single set of two objectives placed at the peaks of the TE_1 mode. (B2, green, circle) consisted of the same objectives as in B1 with the addition of an identical set placed 0.4 μm after the first. This was done to ensure that one objective set was not simply fulfilled locally due to field beating, but that an actual mode conversion was being maintained throughout the waveguide. Finally B3 (dark purple, x) is as B2 with the addition of a third objective put at the node of the TE_1 function implemented to ensure a minimization of the field at this point. In figure 5.2.4(a), the objectives are indicated in different sizes to illustrate the overlapping position, however all objectives were 16 nm^2 corresponding to a single pixel. The flux profile from the resulting simulations are shown in figure 5.2.4(c), where it is apparent that the most even distribution is obtained from configuration B2. It might seem intuitive that increasing the number of objectives would give the designer more control and thus lead to a result closer to that which is sought after, however it also requires that the designer has more knowledge of the necessary weighing of the individual objectives and in turn gives the code less freedom. Here it is found that fewer objectives may in some cases be beneficial for the result of the TO in addition to lowering the calculation time.

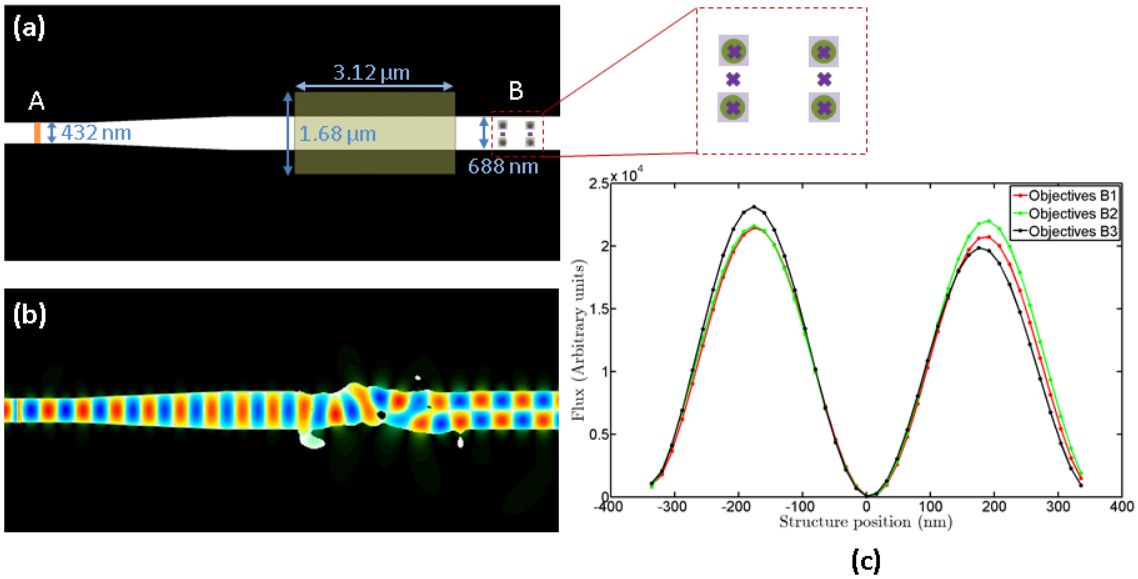


Figure 5.2.4: (a) Starting point structure for the optimization of a straight TE_0-TE_1 mode converter. (b) Mode converter design obtained by applying 2D topology optimization to the structure shown in (a). (c) Comparison of the flux obtained from designs obtained from the three different objective configurations shown with purple in (a).

The design which exhibited the best performance in the simulations was fabricated through EBL, as described in chapter 3, and was subsequently experimentally characterized using the set-up introduced in chapter 4. In figure 5.2.5(a) the SEM

image of the device verifies that design was successfully transferred to the chip. The recorded mode profiles of the device, figure 5.2.5(b), do show functionality and conversion does take place, however it is clear that the device is subject to some beating and that the bandwidth is not as high as sought. The recorded spectrum for the device, figure 5.2.5(c), shows a stable performance, however with a high insertion loss of around 3 dB. This is substantially higher than what the simulations indicate, in spite of the good fabrication results.

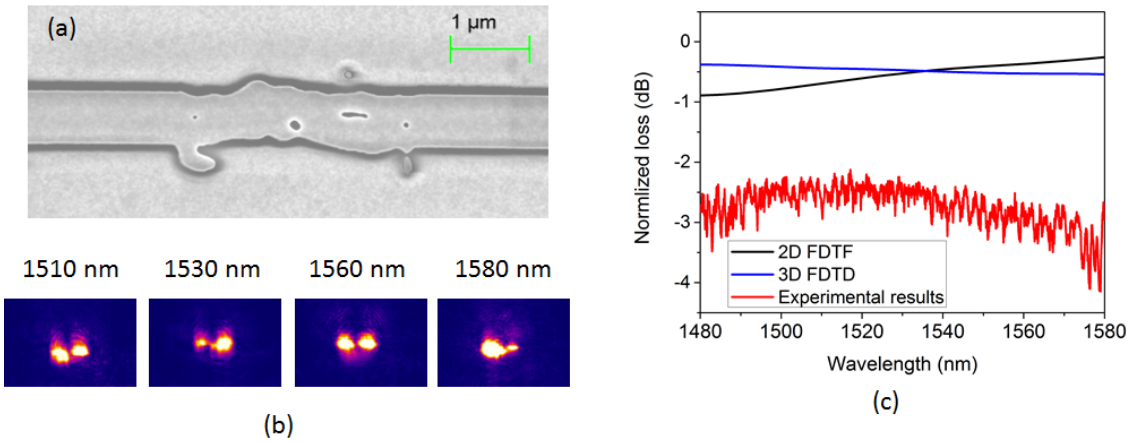


Figure 5.2.5: (a) SEM image of the fabricated straight mode converter. (b) Recordings of the mode profiles of the converted signal at various wavelengths. (c) Theoretical and experimental spectrum for the mode converter. Note that the experimental data has been halved to show the loss of a single device making it comparable to the simulation.

The work presented in the previous section has concluded the necessity of 3D optimization in spite of effective index approximations being applied. Here 2D optimization was attempted none the less in the hopes that the broadband functionality of the device would lead to it working in 3D also. However it is expected that the degradation of the experimental results compared to the simulations originate from this assumption. Further investigation was commenced on the mode profile of the converted signal for simulations done in 3D. The results, shown in figure 5.2.6(a), make it clear that the conversion performance is highly dependent on the wavelength of operation, this in spite of the relatively low loss from the simulated 3D spectrum (blue in figure 5.2.5(c)).

This underlines the hypothesis that the poor performance of the fabricated device originates from the optimization having been performed in 2D. However the work is not irrelevant as it can be used as a starting point for the subsequent 3D optimization. Unfortunately the 3D optimization yielded no well-functioning designs as is clear from the best case device illustrated in figure 5.2.6(b) and (c). The attempts of making a straight photonic waveguide mode converter were thus abandoned.

Square mode converter on straight waveguide

Looking at the 3D design of the straight waveguide, 5.2.6(b), it was hypothesized that there simply was not enough space for the conversion to take place. Hence another attempt was made where the height of the DD was increased and a square region was inserted in the straight waveguide. The resulting design is shown in figure 5.2.7; it is clearly a great improvement in comparison to the straight waveguide. The result does however still exhibit strong beating, and this is not an ideal design. Numerous different DD dimensions were attempted, but even for quite large ones, ranging all the way to $4\,\mu\text{m} \times 5\,\mu\text{m}$ did not yield a satisfactory performance.

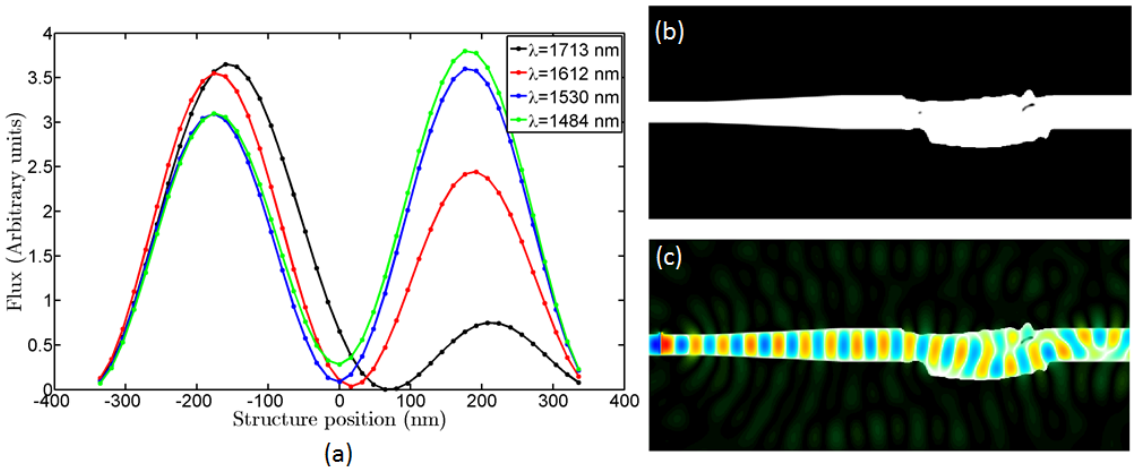


Figure 5.2.6: (a) The flux, simulated in 3D, through the 2D optimized structure, compared for various wavelengths of operation. (b) Design obtained from attempted 3D optimization on the starting point structure of figure 5.2.4(a). (c) The design of (b) overlaid with a H_z field simulation.

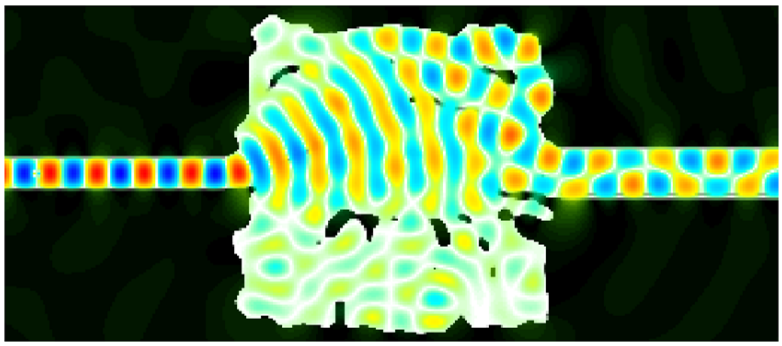


Figure 5.2.7: Topology optimized design of a TE_0-to-TE_1 mode converter made in a square $4\,\mu m \times 4\,\mu m$ region on a straight waveguide presented along with a H_z field simulation.

5.2.3 S-bend structure

As the design of the straight mode converter, described in the previous section, did not seem possible we looked elsewhere for inspiration. In the field of microwave technology, which of course works on a different scale, a compact TE_0-to-TE_1 mode converter has been realized [40], see figure 5.2.8. The quite simple dual-bend, or s-bend, geometry is enough to cause the mode conversion given a correct choice of parameters.

It was attempted to adapt this method for our mode converter centred at 1550 nm rather than the 35 mm wavelength of the microwave system. In equation (8) of [40] the conversion efficiency of the s-bend is stated, and from there it is deduced that the maximum conversion efficiency is achieved when

$$\alpha^2 = \xi^2 R^2\,, \text{ and} \tag{23}$$

$$\frac{\pi}{2} = \theta\sqrt{\alpha^2 + \xi^2 R^2} \tag{24}$$

where we have taken $R_1 = -R_2 = R$ and similarly $\theta_1 = -\theta_2 = \theta$ for the symmetric s-bend in accordance with figure 5.2.8(a). Here the parameters of α and ξ are defined as:

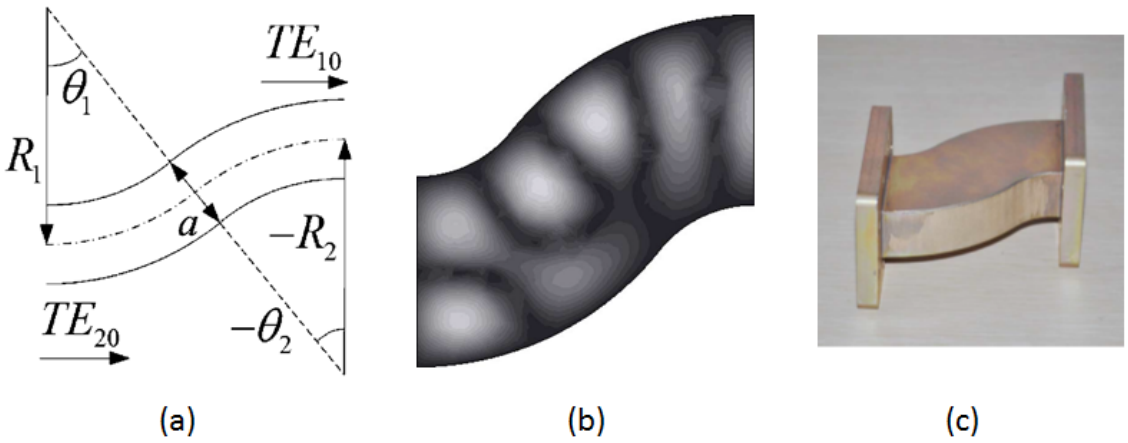


Figure 5.2.8: Figures from reference [40]. (a) Geometry sketch of the basis for the dual-bend mode converter. (b) Simulation of the electric field distribution through the dual-bend structure. (c) The fabricated mode converter device. The size of the structure is in the rang eof 10 *cm.*

$$\alpha = \frac{4}{9}\frac{\beta_{TE_0}\beta_{TE_1}2a^2 + 2k^2a^2 - 5\pi^2}{a\pi^2\sqrt{\beta_{TE_0}\beta_{TE_1}}} \tag{25}$$

$$\xi = \frac{\beta_{TE_1} - \beta_{TE_0}}{2} \tag{26}$$

wherein a is the cross-section of the waveguide, k is the wavenumber, and $\beta = kn$ is the propagation constant. ModeSolutions was used to numerically determine the effective refractive indices of the modes, they were found to be $n_{TE_0} = 2.885472$ and $n_{TE_1} = 2.062466$. Unfortunately it seems that the model breaks down possibly due to the necessity of the fixed thickness of the SOI structure, as calculations leads to: $R = -232.4\,\text{nm}$ and $\theta = 164\,°$.

In spite of the discouraging results of the calculations, a few different bend radii were attempted, to see if it would be possible to gain conversion like this even if suboptimal. In figure 5.2.9(a) an example structure is shown along with the simulated field. Similar to the straight mode converter, the structure has two different waveguide sizes, allowing for the fundamental mode to be injected in a narrow waveguide which will support only that. The waveguide is then gently tapered out to a width that will allow for the TE_1 mode also before the dual-bend occurs. The illustrated example has a bend angle of 46° which was found to be the best performing angle. It is clear that conversion takes place, however substantial beating is present. A number of different bend angles were investigated to map out the effect, the resulting flux simulations are presented in figure 5.2.9. When altering the angle the position of the waveguides were altered to accommodate the changes. None of the angles give an ideal, even distribution of the flux, but the 46° (cyan) bend gives the closest option, wherefore it was fabricated for experimental characterization.

An SEM image of the fabricated structure is shown in figure 5.2.10 along with the data collected for the device. Although the spectrum shows quite a low loss it is somewhat resonant and it is clear from the recorded mode profiles that the conversion neither has a good extinction ratio, nor does it occur for the full bandwidth.

The investigation of the s-bend did not lead to a good design in itself, however it is clear that the technique holds some merit. It illuminates the fact that the positions

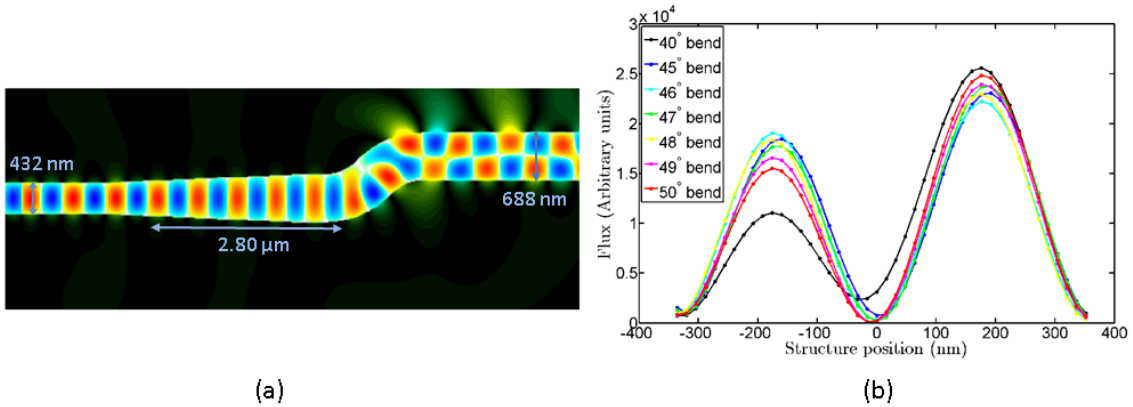


Figure 5.2.9: (a) The s-bend design with a 46° dual-bend is shown overlaid with a H_z-field simulation. (b) Comparison of the flux simulated through s-bend structures with differing bend angles.

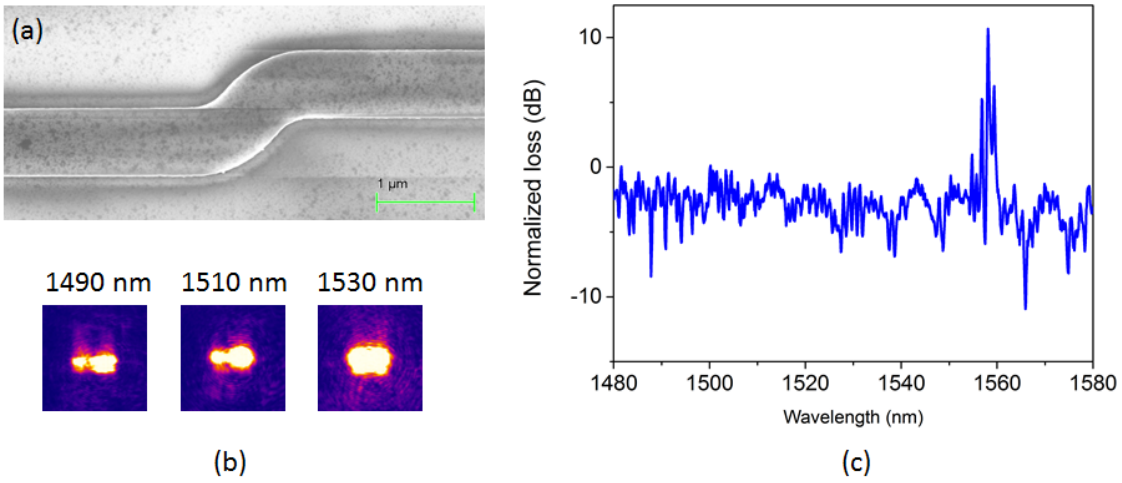


Figure 5.2.10: (a) SEM image of the fabricated s-bend structure. (b) IR-cam recordings of the mode profiles of the signal converted through the s-bend at various wavelengths. (c) Recording of the transmission spectrum of two mirrored s-bend structures, the peak near 1560 nm is caused by system noise, and should be disregarded.

of the waveguide has a big importance for this device; this will be investigated in the following section. It should however be mentioned here that in 2014 Guan et al. demonstrated the theoretical possibility of a lossless (99.5 % conversion efficiency) s-bend mode converter [41]. The key difference seems to be the size, as they achieve this conversion with a fixed vertical offset of 1.2 μm and a length of the bend region of 6.3 μm in comparison to our attempted $\approx 2\,\mu$m.

5.2.4 Optimized bend mode converter

Based on the previous results of this section it is clear that it is necessary to do a full 3D optimization to obtain an efficient and compact mode converter. The attempts with the s-bend structure did not lead to a good conversion design in itself, but it did make it clear that the waveguide position is of great importance. The next step, presented here, is therefore to use a combined starting point structure where TO is applied to a square region over what would be the bend region of an s-bend. Such a starting point is shown in figure 5.2.11(a).

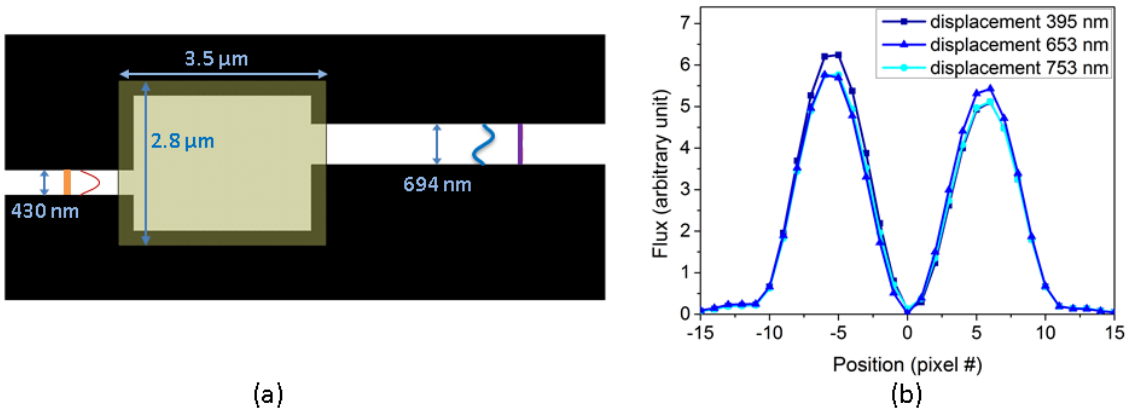


Figure 5.2.11: (a) Starting point structure for the bend topology optimized mode converter, with a displacement of 653 nm between input and output waveguide. (b) Simulated flux of three different configurations of the input and output arms.

Various attempts were made using different sizes of the DD and displacement distances for the input and output waveguides. In figure 5.2.11(b), the simulated flux is shown for three different of these configurations. It is evident that a wrong positioning of the waveguides leads to a poorer design with more beating. Thus having found a proper positioning of the PhWs, it was also investigated which type of objective would lead to the best results. A functional design is obtained in either case, as can be seen in figure 5.2.12(a), however it is clear from the spectra, figure 5.2.12(b), that the performance is worse for the PEO. This is slightly surprising given that the PEO is the simpler one, and that it would thus be more difficult to make human mistakes in the set-up. The relatively simple design, should also be very well suited for the PEO. It is possible that the PSO is more suited for obtaining the desired broadband functionality. It is also quite possible that forcing the slightly longer time-delay allows the optimization to find a better suited local optimum. Whatever the cause, the conclusion is clear from the spectra, and because of that we use the PSO for all future optimizations of the mode converter.

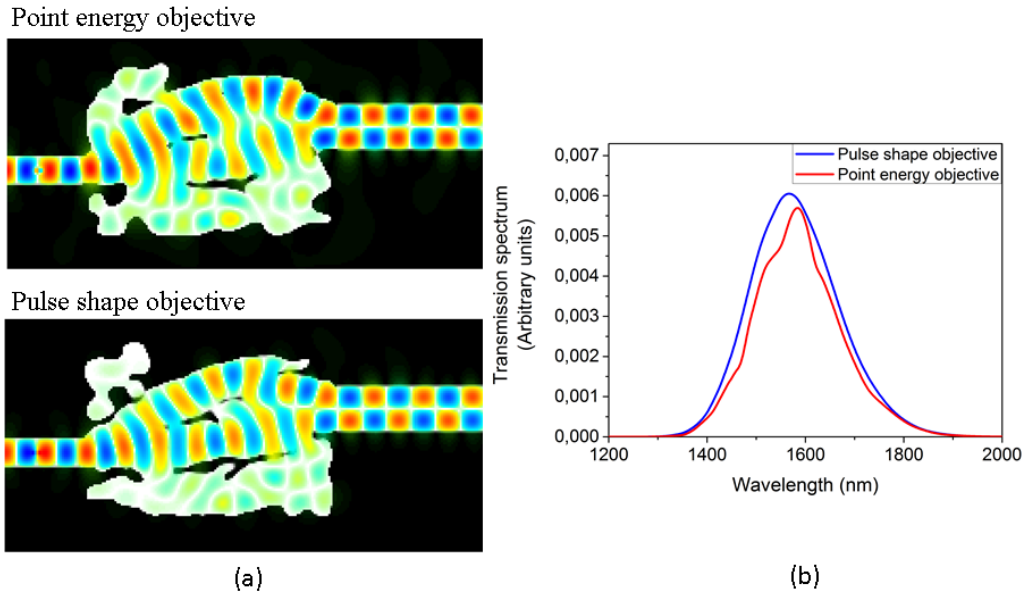


Figure 5.2.12: Comparison of the two objective types used on the same starting point structure and optimizing for the same device type. (a) The H_z field simulations through the structures and (b) the raw spectra of the light transmitted through the mode converter.

In both of the designs it is however quite clear what the optimization is doing. It creates a sort of very compact interferometer where a π phase shift is introduced in one arm, very similar to the principle used in [33] but on a smaller scale. There is a noticeable amount of the field leaking into the side structures and they do take part in lowering the loss, however the main principle is quite clear and intuitive from these structures.

The well-functioning design shown in the top of 5.2.12(a) was fabricated and measured. An SEM image of the final structure can be seen in figure 5.2.13(a); although even the small features of the design were transferred to the chip without any problems, some uncharacteristic side-wall roughness and dirt is observed. This roughness is expected to be the main cause of the losses found for the experimental results in figure 5.2.13(c) being substantially higher than expected from the simulations. However some of this discrepancy might also rise from the long travelling distance of 99.4 μm which the TE_1 mode traverses before the signal is converted back to the fundamental mode on the chip. Regardless of these loss-inducing circumstances the device exhibits a 3 dB-bandwidth of the entire 100 nm bandwidth of the source. The qualitative functionality of the device is apparent from the mode profiles shown in figure 5.2.13(b) and (c).

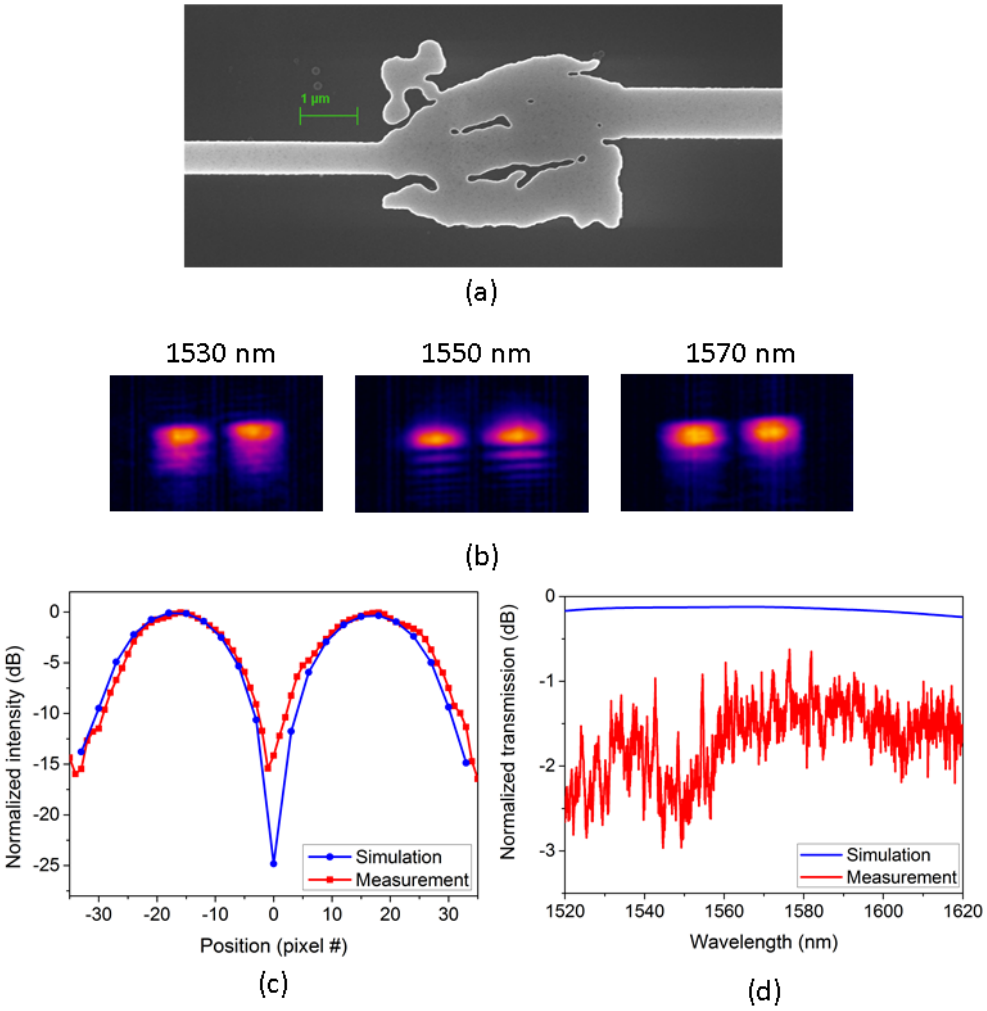


Figure 5.2.13: Experimental results for the 3D topology optimized mode converter design with a waveguide displacement. (a) SEM image of the fabricated structure. (b) IR recordings of the converted signal. (c) Cross-section of the mode profile at 1570 *nm, both simulated (blue) and experimental (red). (d) Recorded and simulated transmission spectra for the device. Note that the experimental data (red) has been halved from the measurement, which included two devices spaced by* 99.4 μ*m of a straight waveguide carrying the* TE_0 *signal.*

5.2.5 Re-optimization of the mode converter

It has now been demonstrated that it is possible to make a compact mode converter from TE_0-to-TE_1 using TO. It functions using a relatively simple interferometer scheme, and this spurred the investigation of making an even more compact device with a slightly more refined starting point structure, figure 5.2.14(a). Here the knowledge of the design from previous optimizations is used, and a slightly elongated square silicon region with a hole in the center is the input for the optimization. As illustrated in figure 5.2.14(b), it is far from enough to create the conversion in itself, however a better starting point, will guide the optimization closer to a favourable local optimum from the beginning and will thus speed up the optimization as well as guide it in the desired direction.

Figure 5.2.14(c) and (d) show the topology optimized design along with the corresponding field. Although the design is smaller than that shown in the previous section and it does not have the same side-structures, there are some prominent similarities. Both designs have a small dot of air before the longer split between the waveguide arms in the interferometer structure, and both structures introduce a small side-structure in the upper left corner of the DD as well as an expansion of the waveguide right before the end of the DD. These features are the ones that allow the conversion on such a small scale, and which distinguish the design from he basic principles of the interferometer.

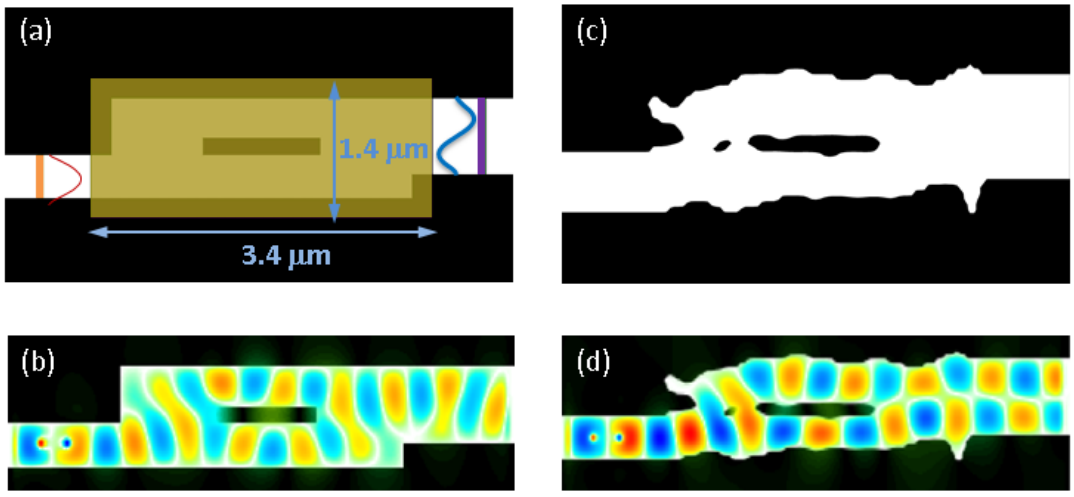


Figure 5.2.14: (a) Starting point structure for the optimization of a TE_0-to-TE_1 mode converter based on the knowledge from previous designs shown in figure 5.2.12. (b) H_z field simulated through the starting point structure. No conversion is taking place. (c) The topology optimized structure on its own and (d) along with the simulated H_z field.

In figure 5.2.15 the performance, both theoretical and experimental, of the re-optimized structure, using the more refined input, is presented. The conversion is found to take place across the entire range of the tunable laser source, however beating is present. In figure 5.2.15(b) it is clear that the extinction ratio of the mode profile is resolution limited by the IR camera. For the more compact converter it reaches $> 9\,\mathrm{dB}$. It is more remarkable that the structure exhibits even lower losses, $< 1\,\mathrm{dB}$ for the majority of the range of the laser source, than the corresponding larger device described in the previous section. This discrepancy is likely to be due to the poorer quality of the fabricated chip for the previous device. A comparison of the simulated spectra clearly shows that the larger device should perform better. The other major difference between the performance of these two devices is the bandwidth of the structure.

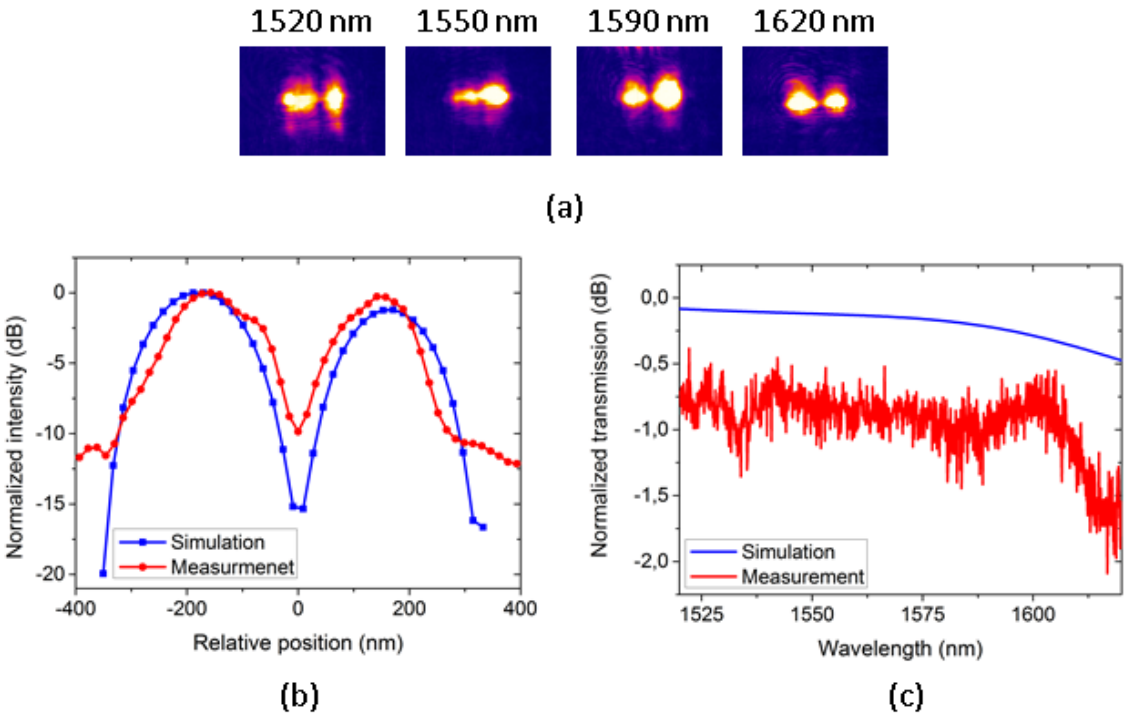


Figure 5.2.15: (a) Recorded mode profiles of the compact TE_0-to-TE_1 mode convert, the design of which is shown in figure 5.2.14(c). (b) Theoretical and experimental line scans across the mode profile at 1600 nm. (c) The recorded transmission spectrum for a single device along with the simulated data.

Looking at a larger range of the simulated spectrum in figure 5.2.16(b), it becomes clear that it is mainly the bandwidth which is affected by the change in footprint. The bandwidth is shifted between the two cases, but also generally narrowed a bit for the smaller design, where the transmission spectrum drops off at higher wavelengths. The slight degradation of the performance does also appear in the simulated mode profiles in figure 5.2.16(a), where the extinction ratio is higher for the bigger design and the peaks of the modes are more evenly distributed as well.

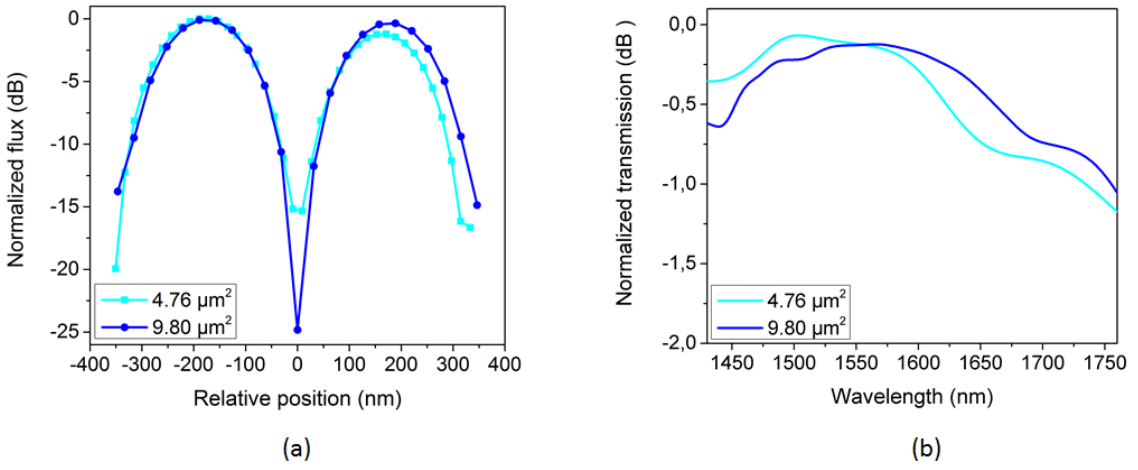


Figure 5.2.16: Comparison of the theoretical performance of the two 3D mode converter designs shown in figure 5.2.11(a) lower part (dark blue) and 5.2.14(c) (light blue). (a) Normalized flux and (b) simulated transmission spectra.

5.2.6 Higher order mode converter

The previous sections have concerned the importance of the starting point structure and have illustrated how fine tuning of those can lead to even better performances. Here the claim that, knowing just the basics of a simple starting point, TO can be used for increased complexity designs. With offset in the knowledge of the work with the TE_0-to-TE_1 mode converter a higher order TE_0-to-TE_2 mode converter is designed.

The basic principle is identical to that presented in section 5.2.4, but a few minor adjustments are made. The starting point structure for the higher order mode converter is shown in figure 5.2.17. To enable the presence of the TE_2 mode the width of the higher order mode waveguide is increased to 1.24 μm. The size of the DD is also increased as is the height difference of the two waveguides. Three different DD configurations are tested: Two are square with the difference that one is made large enough to include the connecting region of the waveguides and the last ignores the corners of the square silicon region as they are expected to be of little relevance to the design.

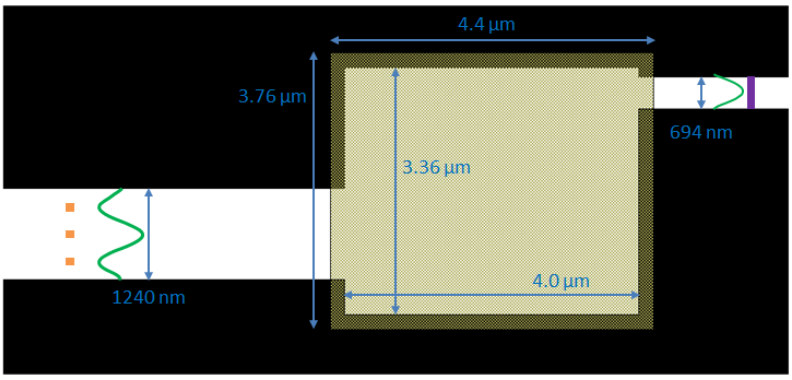


Figure 5.2.17: Starting point structure for the design of a TE_0-to-TE_2 mode converter.

In figure 5.2.18 the topology optimized designs obtained from the different DDs and with three different filter sizes are presented. The overall features of the structures are the same even as the filter size is increased and all of the structures exhibit the desired functionality. The performance is more directly compared in figure 5.2.19. From the spectra it is evident that most of the square region devices have very similar functionality exhibiting low losses < 1 dB. The 3 dB-bandwidth of around 110 nm is comparatively low and is shifted towards higher wavelengths. The loss of the 80 nm filter is only marginally different from the smaller 40 nm filter. The largest filter and the cut off structure does however exhibit significantly increased losses.

Although the design could be tuned and modified further, as has been done for the lower order mode case, this design has demonstrated the ease with which the TO designs can be expanded to include more complex functionality.

The structure was attempted fabricated, however some complications were met, as it was not possible to use the standardized fabrication method described in chapter 3. As the ASE was out of use at the time of fabrication a different machine was used for the etching, namely the Advanced Deep Reactive Ion Etch (DRIE) utilizing Pegasus ASE, MP0636[2]. Furthermore a different resist, CSAR, was used for the etch mask. This process has not been adjusted and fine-tuned for the fabrication of these small

[2]Many thanks to Hitesh Sahoo Kumar and Xiaowei Guan for helping with the fabrication using this, to me unknown, equipment.

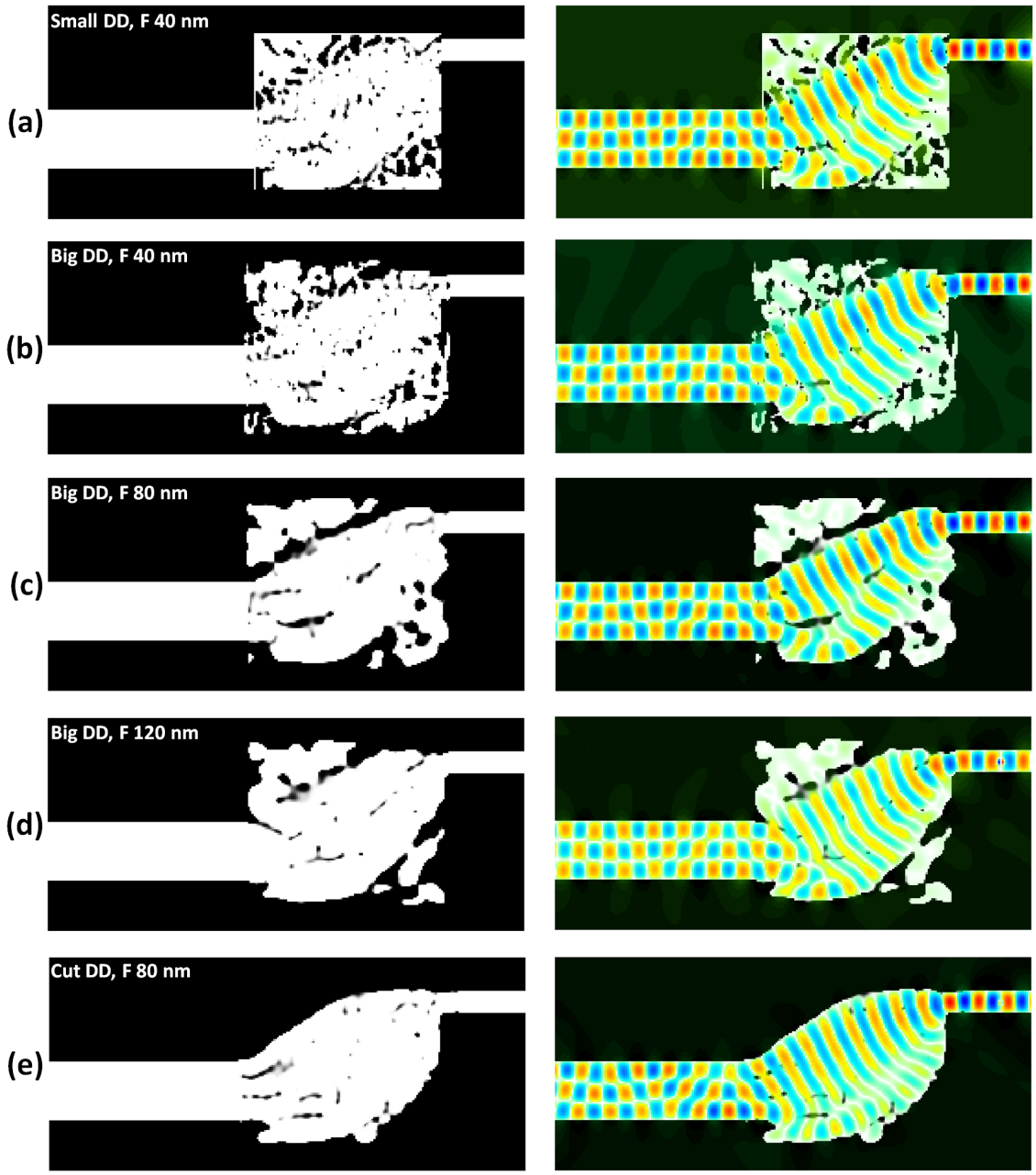


Figure 5.2.18: Left column: The topology optimized TE_0-to-TE_2 mode converter designs for different optimization settings. Right column: The same structures with the 3D FDTD calculated H_z-field overlaid. (a) DD has the same size as the square Si slab. (b) DD extends beyond the Si slab to include the waveguide connections, filter 40 nm. (c) The large DD with a filter of 80 nm. (d) Filter size increased to 120 nm and another 100 iterations run to remove grey areas, for the same large DD structure. (e) Corners of the DD and the underlying Si slab removed, filter size 80 nm

feature structures. An SEM image of the fabricated structure is shown in figure 5.2.20(a). Although the overall structure has been transferred to the chip, some of the smaller features have disappeared during processing. Furthermore areas of a different shade of grey can be noticed around the edges of the structure indicating residual material - either because resist remains on the chip or because the etch was not complete. As is evident from the recorded mode profiles shown in figure 5.2.20(c), these complications during fabrication have led to the device not being functional. For the high wavelength case there is a hint of the TE_2 mode, however it is noisy and not maintained at other wavelengths. It is however believed that given

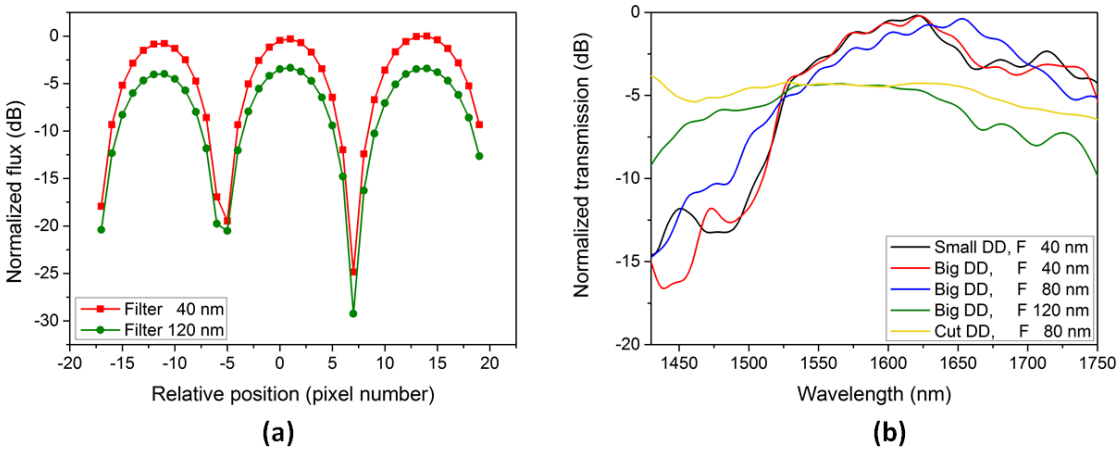


Figure 5.2.19: (a) Simulated flux of the TE_2 mode after conversion in the optimized devices with smallest (40 nm) and largest (120 nm) filter sizes, both for the larger design domain. (b) Comparison of the transmission spectra for the optimized designs presented in figure 5.2.18.

the opportunity to use the regular equipment for fabrication, or refining the process parameters for the DRIE etch and CSAR resist, it will be possible to fabricate the mode converter retaining both functionality and low loss.

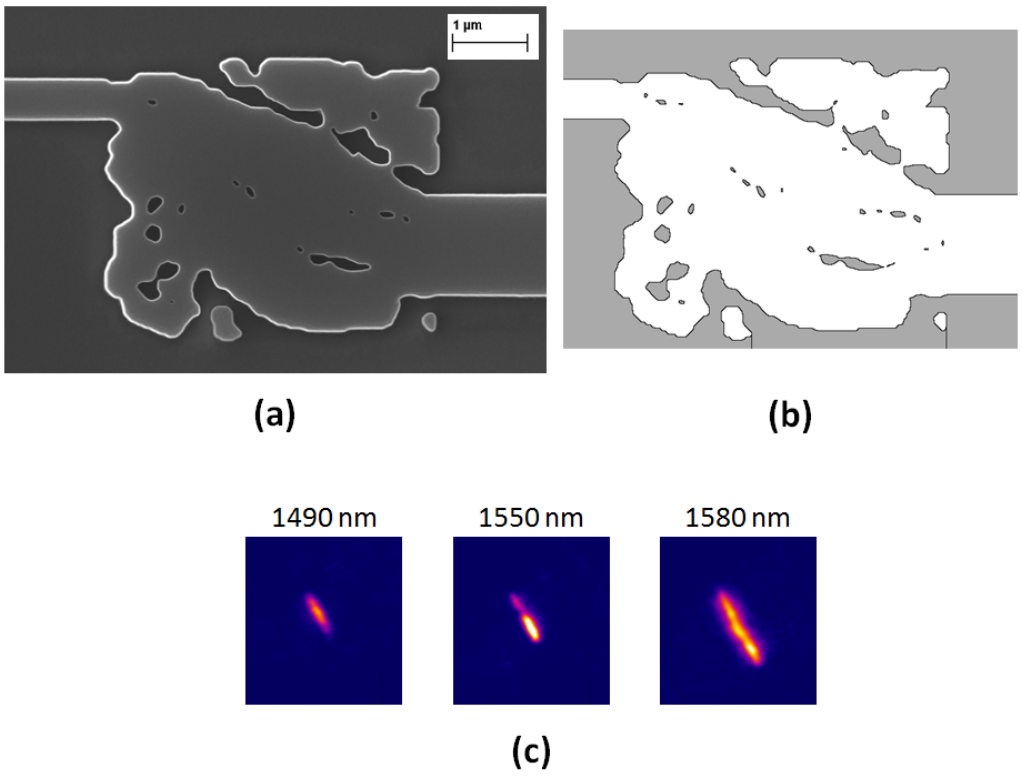


Figure 5.2.20: (a) SEM image of the TE_0-to-TE_2 mode converter designed with an 80 nm filter and fabricated with a non-routine process flow. (b) Image from the design file, showing the design intended for writing on the chip for easy comparison. (c) Measured mode profiles after intended conversion to the TE_2 mode recorded at three different wavelength. Unfortunately the mode has clearly not be converted as desired.

5.3 2-mode Multiplexer

The mode converters described in the previous section are naturally important for the sake of controlling and manipulating the modes on-chip. However the main reason for the interest in working with higher order modes is the possibility of multiplexing and thus transferring more data simultaneously. This is relevant on and off chip, where potentially a multiplexed fibre signal could be coupled to the chip directly and then be compactly converted in the PIC. In this section the design of a TE_0 and TE_1 mode (de-)multiplexer is described.

Typically integrated MDMs have been relying on asymmetric directional couplers [42] [43], Y-junctions [44, 45, 46, 47], or multimode interferometers [48, 49, 50]. Below the working principles as well as advantages and disadvantages are briefly outlined.

Asymmetric directional coupling multiplexers are based on fulfilling the conditions of phase matching. Two waveguides are made to run parallel and in the active region they are brought close while their widths are adjusted to ensure that the effective refractive indices are matched for the modes that are to be multiplexed [42]. A simple illustration of the concept for a two mode TE_0-to-TE_1 (de-)multiplexer is shown in figure 5.3.1. The TE_0 mode is input on the left in both waveguides. One of them is only wide enough to carry the TE_0 mode, the other wide enough to carry both TE_0 and TE_1. The widths have been carefully chosen to ensure fulfilment of the phase matching condition $n_{eff,0} = n_{eff,1}$, meaning that the effective index of the fundamental mode carried in the narrow waveguide is the same as the effective index of the TE_1 mode when propagating in the wide waveguide. In the coupling region the two waveguides are brought close and the light of the narrow waveguide is injected into the wider one, leading to a multiplexed signal exiting to the right.

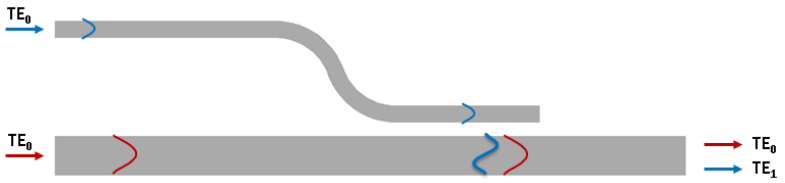


Figure 5.3.1: Sketch of the working principle of an asymmetric directional coupler based mode (de-)multiplexer operating on the TE_0 and TE_1 modes. By ensuring that the effective refractive index of a narrow waveguide supporting only the $TE_0$1 mode is the same as the effective refractive index of a TE_1 mode of a wider waveguide the modes can be coupled selectively leading to multiplexing.

This method has a high coupling efficiency leading to good conversion and very low mode crosstalk, furthermore it can operate with a high bandwidth [43]. However due to the need of satisfying the phase matching condition, which puts constraints on the waveguide widths, lengths and gaps, the devices are quite sensitive to fabrication errors - methods to lax these constraints are however being investigated [42]. These devices tend to have lengths of $\sim 50\,\mathrm{nm}$. An increased number of mods can be obtained through cascading of such devices where the wide waveguide is tapered to maintain phase-matching.

Ring-resonator based MDMs have also been employed in different variations [51][52], one of which operates quite similarly to the asymmetric directional couplers. The main difference is that coupling does not happen directly from one waveguide

to another, instead light is evanescently coupled from a narrow single mode carrying waveguide to the ring resonator and from that onto a wider bus-waveguide which can carry the multiplexed signal, illustrated in figure 5.3.2(a). The main benefits of using the ring resonator is that it is possible to achieve a much shorter coupling length $< 8\,\mu$m in comparison to the $\sim 50\,\mu$m for the asymmetric directional coupler. The structures can be scaled in a similar manner to include more modes by cascading. Alternatively it is possible to use multimode ring resonators for (de-)multiplexing as the ring can pick one signal out if it is only resonant with either of the modes. The principle is sketched for a TE_0-TE_1 system in figure 5.3.2.

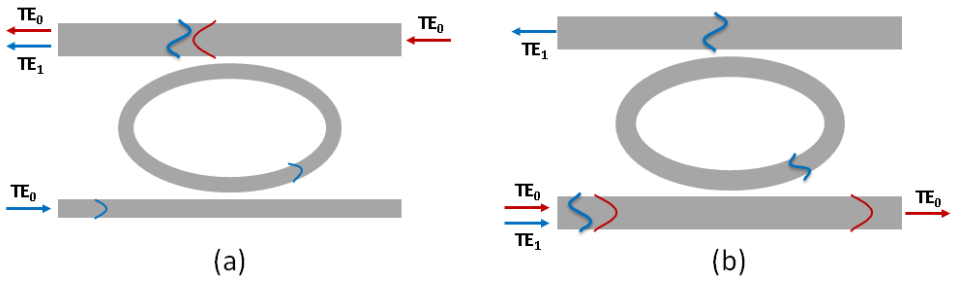


Figure 5.3.2: (a) Sketch of the working principle of a ring resonator mode multiplexer based on coupling to a bus-waveguide. (b) Illustration of a multimode ring resonator based mode multiplexer.

Asymmetric Y-junctions are another method which utilizes the difference in effective refractive index obtained from varying the width of the waveguide. The mode-splitting behaviour of these junctions are caused by the effects of mode localization and adiabatic propagation [53]. In the case of the asymmetric Y-junction, sketched in figure 5.3.3, the stem of the Y can carry the multiplexed signal being wide enough to allow for both the TE_0 and the TE_1 modes. Both of the branches are narrow enough that they carry only the fundamental mode. The mode running from the stem will couple to the branch which has an effective refractive index closest to its own. This means that the fundamental mode of the stem will couple into the wide branch which has the closest effective index itself. However the effective index of the narrow branch is closer to that of the TE_1 mode of the stem and thus the splitting takes place [45]. This process is theoretically lossless and reciprocal making it a great mode multiplexer. The power splitting between the two arms can be tuned by carefully designing the waveguide widths and the devices can be made to have very broadband functionality with low insertion loss and cross-talk [44]. The main drawbacks of this method is the rigid requirements for the design in regards to junction angle and waveguide widths [47]. Furthermore the structure is quite sensitive to fabrication errors for the sharp junction and the induction of polarization rotation [46]. These devices have lengths of up to a few hundred nm.

The method can be extended to include more modes by widening the stem and increasing the number of branches of the junction. The relative phase difference accumulated traversing the junction will influence the output power but the few-mode asymmetric Y-junctions are approximately independent of the phases as long as they are practically adiabatic and asymmetric enough to ensure that the modes have a clear preference for specific output arms. [53],[45].

Interferrometric multiplexers are based on multimode interference which in turn builds on the self-imaging property. A region of specific length is dedicated

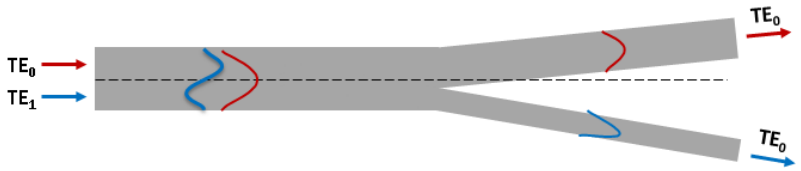


Figure 5.3.3: Sketch of the working principle of asymmetric Y-junction. The wide stem carries the multiplexed signal of the TE_0 and TE_1 modes which are de-multiplexed into the two branches of the Y.

to the constructive interference of the excited fields. The dimensions of the multimode interference region must be properly designed in accordance with the beat length, $L_\pi \cong \frac{4n_c W_{eff}^2}{3\lambda}$, where n_c is the refractive index and W_{eff} is the effective width of the region [49]. Multiplexers operating on different numbers of modes can be obtained by combining various multimode interference devices. Here the 2-mode (de-)multiplexer described in [49] will be related as a relatively simple example although there does exist examples of (de-)multiplexers manipulating more modes, such as [48]. The device consists of three sections sketched in figure 5.3.4(a). A mode converter splitter uses the symmetric interference to map the TE_0 mode to the center of the interference region while the TE_1 mode is mapped to the edges and from here coupled out as TE_0 modes. An example is shown in figure 5.3.4(b). After this a phase shifter is placed delaying the output of one of the TE_0 modes originating from the TE_1 mode, so that they will end up with a phase difference of $\pi/2$. This phase difference is necessary in the final step where a 3-dB coupler, based on general interference, allows the central input light of the original TE_0 mode to pass through to its central output. Meanwhile it maps the two TE_0 modes originating from the TE_1 mode, and which now carry a phase difference of $\pi/2$, into one of the outer ports. In this case the original signal has been demultiplexed.

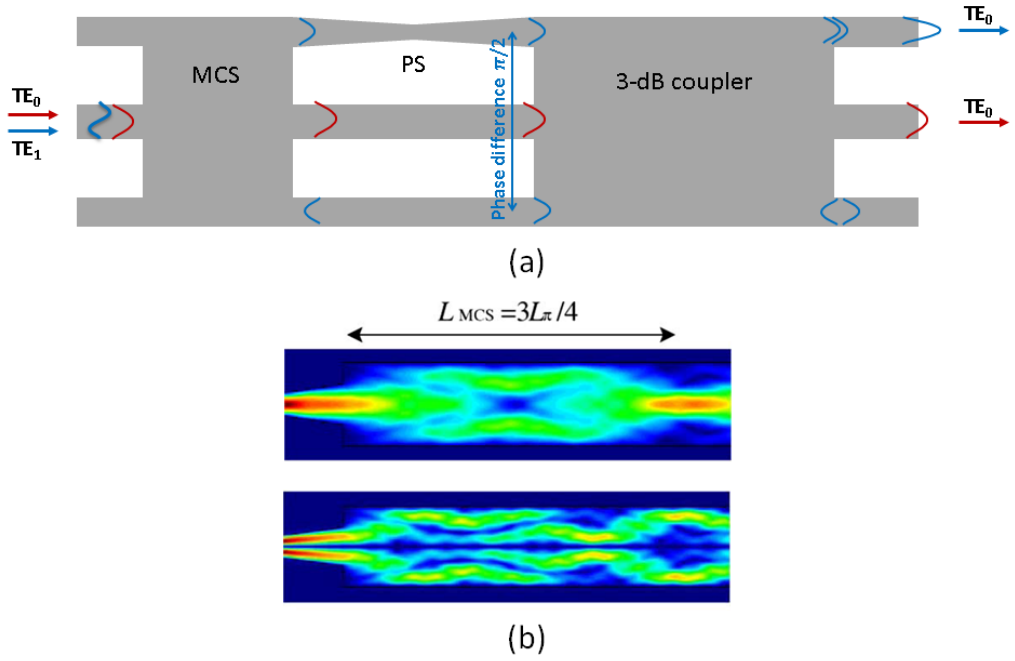


Figure 5.3.4: (a) Sketch of a three-component two-mode multiplexer based on multimode interference. (b) Field distribution of the symmetrically interfering field of a multimode interferometric waveguide for the input of a TE_0 (top) and TE_1 mode (bottom), directly from [49].

A drawback of the interference based multiplexers is that they tend to have somewhat narrower bandwidths than the other methods described here, although still covering the c-band. They do, however, exhibit low loss and they have high tolerance to fabrication errors. A thing that tends to be highlighted is also their relative compactness, the device described here is $< 80\,\mu$m long.

Another interference based device can be found in [50] where a Mach-Zehnder setup is used along with periodic waveguide gratings which allow for the propagation of the second order mode while the first order mode is reflected. 3 dB/90° couplers are included at each end to split/combine the modes.

These are just some examples of working principles of frequently seen solutions for the integrated mode (de-)multiplexer designs, however there does of course exist many more. In this section it will be described how those designed using TO come to be and how well they work.

5.3.1 Initial design attempts

Experience has shown that 3D designs are necessary for good experimental results, nevertheless it was decided to do 2D optimization initially. This was done because of the much lower calculation time of the optimizations in 2D. Furthermore the working assumption was that it would lead to a good basis for optimization parameters which would make the subsequent 3D optimization a much quicker procedure.

As a starting point a symmetric design was made with one input waveguide wide enough to support the TE_1 mode, here 688 nm, and two narrower wavegudies supporting only the TE_0 mode, namely 432 nm. The DD was a square region the size of which was altered for the various designs ranging from $2\,\mu$m to $6\,\mu$m. An example of the structure is shown in figure 5.3.5.

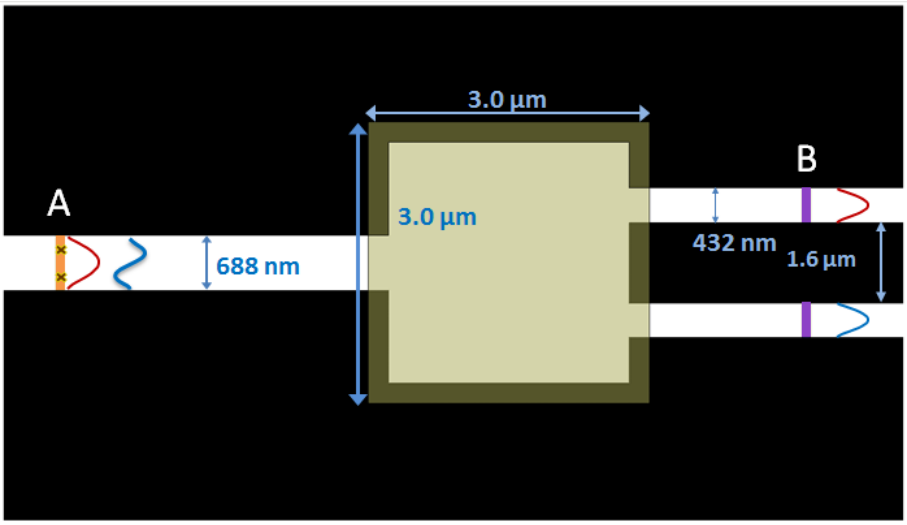


Figure 5.3.5: Example of the symmetric starting point structure for the 2-mode multiplexer working on TE_0 and TE_1. Several variations were made with differing dimensions of the design domain (yellow), however the principle is the same. The orange bar in the left waveguide indicates the source position in the x direction and illustrates the excitation of the fundamental mode. The two brown crosses on top of it indicates the excitation of the first order mode using two point sources with opposite orientation. The objectives of the optimization are placed at the purple bars at position B. The resolution is 16 nm/pixel.

In the optimization of the 2-mode multiplexer, the PSO has been employed. At two different starting times t_{in,TE_0} and $t_{in,TE_1} = t_{in,TE_0} + 400\,\text{fs}$, the TE_0 and TE_1 modes are excited at position A in the broad waveguide. The propagation time of the light through the structure is then estimated though a field simulation and the goal is to maximize the output of the TE_0 mode in the upper waveguide at point B at this arrival time t_{out,TE_0} while minimizing it in the lower waveguide at this same time. Similarly TE_1 is maximized in the lower waveguide at t_{out,TE_1} and minimized in the upper. Here the method of maximizing the output by specifying a low amplitude pulse and optimizing for the pulse shape to be unlike that, as described in 2.1.1 has been employed. This is also the main reason that the design has been made for the demultiplexer configuration rather than that of the multiplexer: It allows for a simpler set of objectives, which gives more control to the TO and leaves less room for the designer to enforce unnecessary limitations. However the designer still has an important task in properly determining the appropriate delay for the arrival at B. A good starting guess is the propagation time of the light traversing the unoptimized structure. However it is necessary to make several iterations where the delay is altered based on the amount of scattering seen in the previous design. The same delay is not ideal for both the TE_0 and the TE_1 mode.

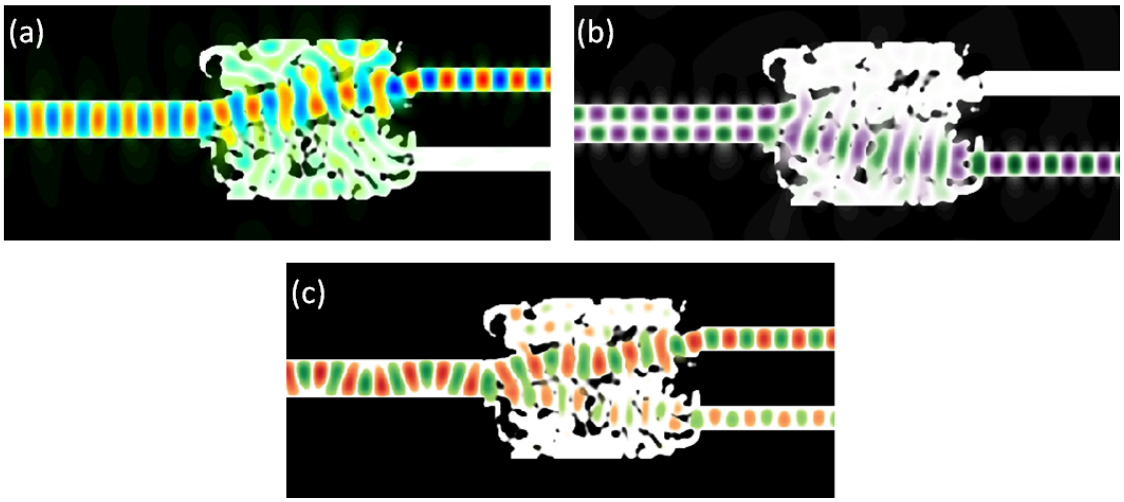


Figure 5.3.6: 2D FDTD simulations of the fields through a 2D optimized 2-mode multiplexer with a footprint of $3\,\mu m \times 3\,\mu m$. (a) The TE_0 field, (b) the TE_1 field, and (c) the two fields excited simultaneously illustrating the multiplexed field which is more difficult to qualitatively evaluate.

A resulting design, obtained from the starting point shown in figure 5.3.5 is given in figure 5.3.6. The field simulations clearly show the functionality verifying that this is a potential 2D design to continue work from. However there is a lot of scattering for the TE_0 mode in spite of it practically only needing to be guided from one waveguide to the other. The next step in the design investigation was thus the position of the waveguide, as this was found to be a very important feature for the mode converters. The hypothesis for this design was that the TE_0 mode needed to be moved only very little. A perfect alignment of input and output would however not be possible, as that would lead to obstructing features necessary for the conversion of the TE_0 to TE_1 mode to get in the way of the ideally unchanged TE_0 mode.

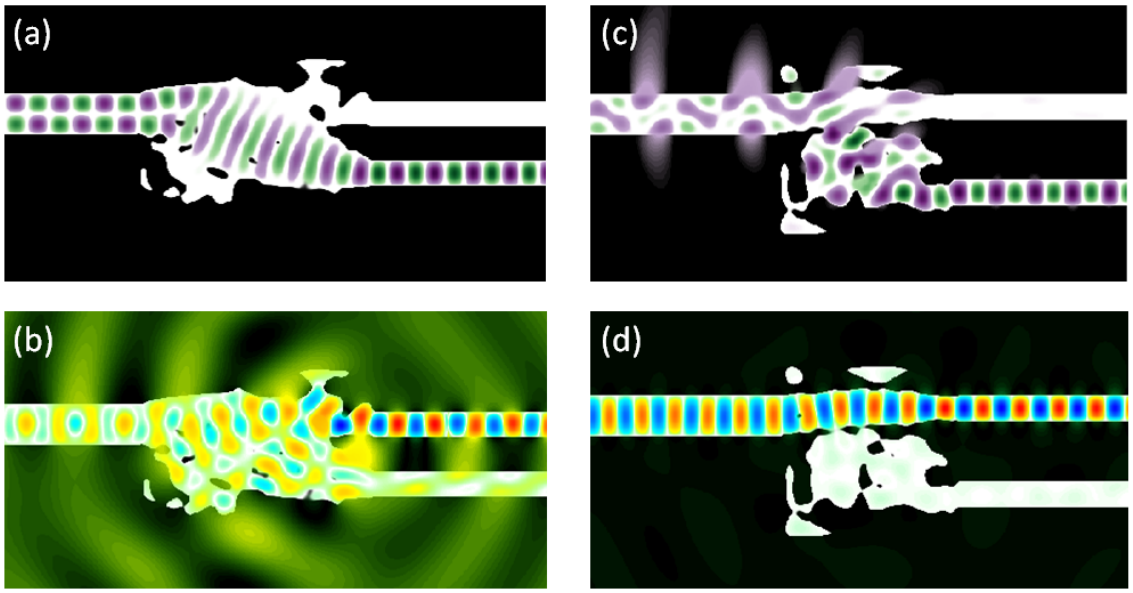


Figure 5.3.7: 2D FDTD simulations of the fields being multiplexed through two different attempted designs of a 2-mode multiplexer. (a) TE_1 and (b) TE_0 through a $2\,\mu m \times 3\,\mu m$ design. (c) TE_1 and (d) TE_0 through a $3\,\mu m \times 3\,\mu m$ design. Note that the size is not the only difference of the starting point, the waveguide positions are also key parameters for the designs. (a) and (d) were the best designs obtained for the two respective channels.

In figure 5.3.7 two of the 43 different design configurations that were attempted are shown along with the 2D FDTD simulations of the fields being multiplexed in the structure. These structures are clear examples of one of the major problems with the designs: They would often function very well for one of the modes but poorly for the other, either not performing the desired conversion or simply incurring great losses. An attempted solution for this problem was to manually combine the two designs. The waveguide displacements were maintained as was the overall outline of the structure. Figure 5.3.8 shows the two designs along with the starting point structure made from them, figure 5.3.8(c). Topology optimization was applied to this new collected design and the results were good enough that further 3D optimization was commenced.

The starting point structure used for this 3D design, was the one shown in figure 5.3.8(c), however objectives as well as time delays needed to be adjusted when going from 2D. Hence the resulting design presented in figure 5.3.8(e) was obtained after several iterations. The simulations of the performance of the optimized designs are given in figure 5.3.9. The 2D and 3D designs have some similarity, yet are quite different. In the transmission spectra of the 3D FDTD for both structure, found in figure 5.3.9(b), it is clear that the difference in the two designs have great impact on the performance. 3D TO has once more been shown to be necessary for good performance in spite of the broadband functionality of the device.

The device was fabricated in accordance with the description in chapter 3 and experimentally characterized with the set-up presented in chapter 4, the results are shown in figure 5.3.10. The recorded mode profiles show the clear functionality of the device across the entire 100 nm bandwidth of the laser source. The extinction ratio of the TE_1 mode was found to be $> 9\,dB$ while the insertion loss of a single device was found to be $< 1.3\,db$ and the cross-talk around $-12\,dB$. All of this was achieved for a footprint of around $4.4\,\mu m \times 2.8\,\mu m$ making it, to the best of our knowledge, the worlds smallest mode multiplexer.

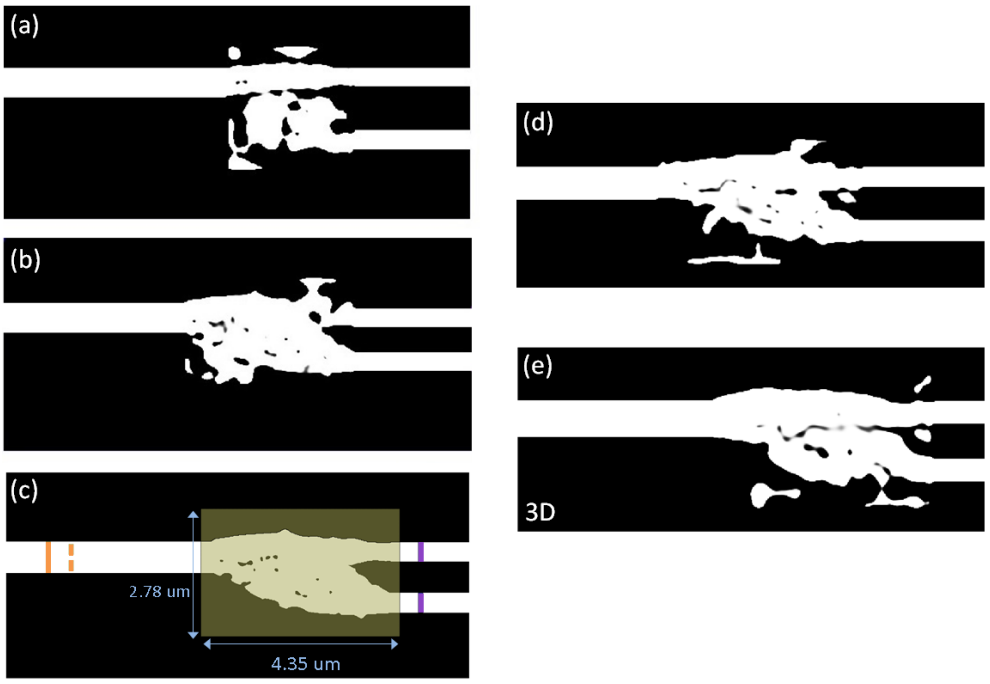


Figure 5.3.8: (a) and (b) Illustrations of two partially functional optimized designs (shown along with field simulations in figure 5.3.7). (c) A starting point structure obtained from manually collecting the two designs shown in (a) and (b). (d) Design resulting from 2D TO performed on the structure given in (c). (e) Design resulting from 3D TO performed on the structure given in (c).

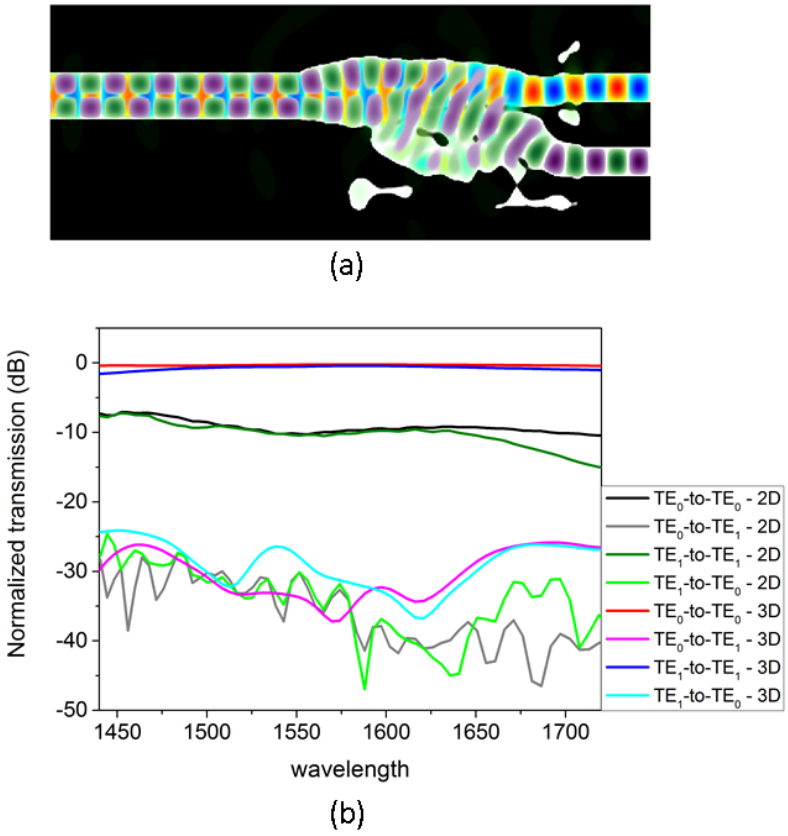


Figure 5.3.9: (a) 3D FDTD simulation of two TE_0 fields being multiplexed. The fields shown are overlays of each field being processed in turn, as the actually multiplexed signal is not easily recognizable (example shown in figure 5.3.6(c)). (b) The simulated transmission spectra of the 3D FDTD calculations through the structures optimized in 2D (5.3.8(d)) and 3D (5.3.8(e)), respectively.

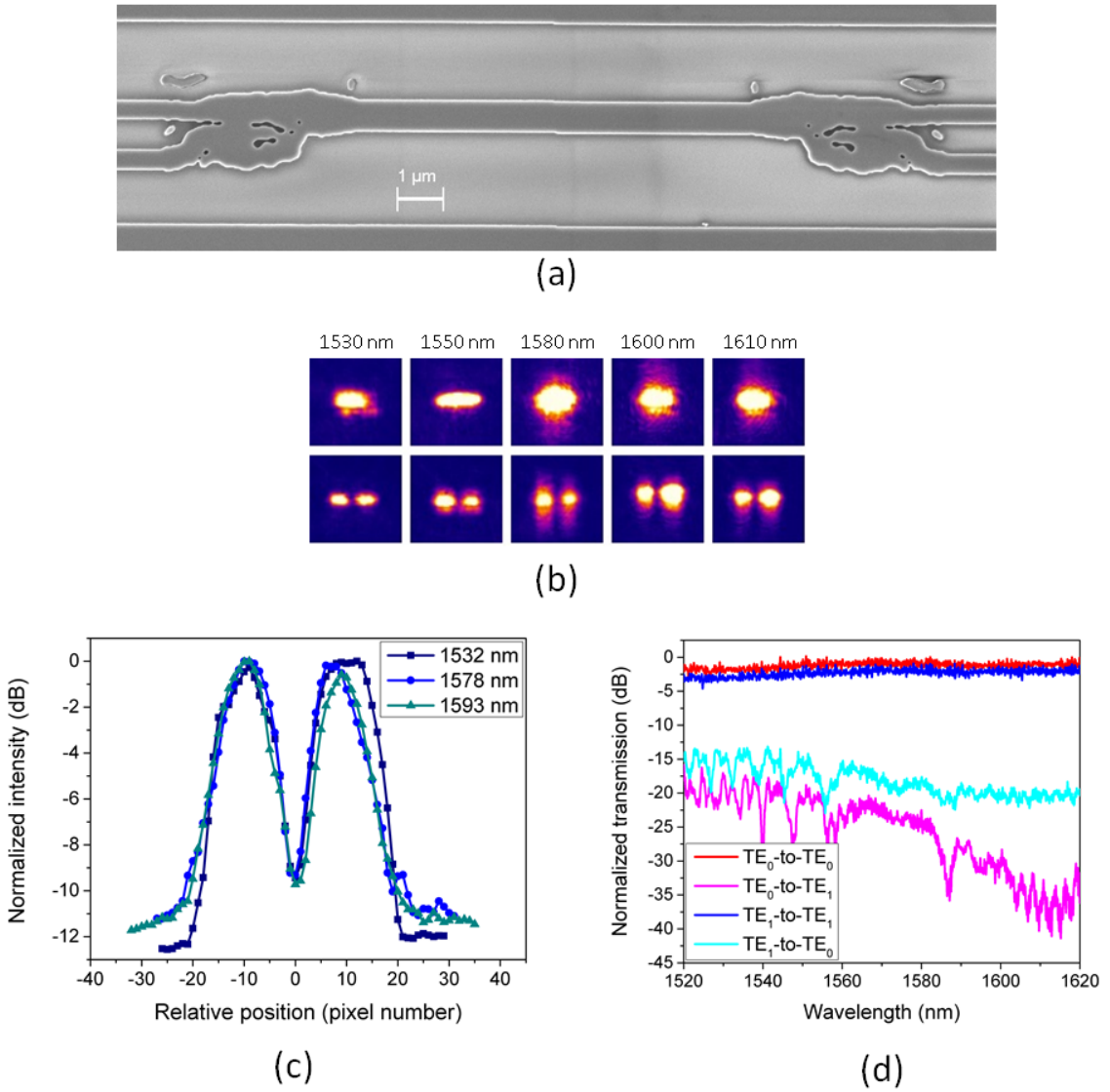


Figure 5.3.10: (a) SEM image of the fabricated double structure of the 3D topology optimized 2-mode multiplexer. (b) Mode profiles recorded at various wavelengths for light input in either of the two TE_0 channels, showing the maintained TE_0 mode and the converted TE_1 mode respectively. (c) A line scan across the recorded TE_1 mode profile for three different wavelengths. (d) Transmission spectra recorded for all four combinations of input and output channels.

5.3.2 Square input structure

Although the mode multiplexer device presented in the previous section was highly functional, further investigation was commenced. Topology optimization played a central role in the device design, however by using the manually combined device shown in figure 5.3.8(c) as a starting point it does not fully illustrate the strength of the arbitrary structure TO. Naturally there are many parameters that the designer changes whereby the outcome is influenced and the starting point structure is only one of them. Choosing a clever starting point has merit in and of itself, however the goal of this section is to illustrate, how it is not a necessity to have such a refined starting point for TO to deliver a functional design.

The new starting point is given in figure 5.3.11(a). The optimization was attempted with waveguide positions maintained from the structure in figure 5.3.8(c), but variations were also made, shifting the waveguide which carries the demultiplexed TE_1 (the lower one) up and down. Here it was found that by shifting the waveguide upwards it was possible to use a slightly smaller design domain, and thus this was the configuration that we proceeded with. Aside from the starting point structure only the propagation time though the multiplexer has been changed. As previous work showed that 3D optimization is necessary for proper experimental functionality and that extensive tuning of the optimization parameters is still necessary when going from 2D to 3D designs, the following was only done in 3D. The resulting design is shown in figure 5.3.11(b) and is noticeably different from the designs presented in the previous section although the method for the optimization is the same.

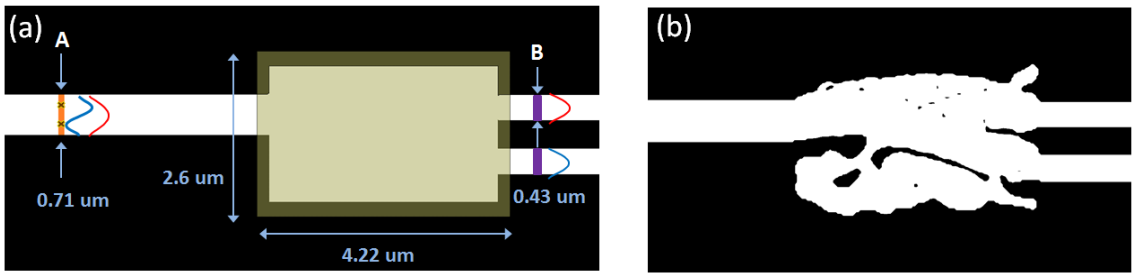


Figure 5.3.11: (a) New starting point structure for a 2-mode multiplexer based on a rectangular region but an asymmetric starting point. (b) The 3D topology optimized design.

The performance of the device was investigated through simulations and the results are shown in figure 5.3.12. Generally the performance of the structure is good, with an insertion loss of $< 1\,\mathrm{dB}$ for the entire 350 nm bandwidth of the source and an extinction ratio of $> 18\,\mathrm{dB}$. The H_z-field simulations have been done at various wavelengths, and it was found that very little of the field was carried in the 'appendix' at the bottom of the structure. The hypothesis was, based on the results of section 5.2.5, that this appendix allowed for the broadband functionality. This theory was however disproved by the simple test of manually removing the appendix from the structure and redoing the simulation, the results are shown in figure 5.3.13. Here it is clear that although removing the appendix generally increases the loss of the (de-)multiplexer it has hardly any impact on the bandwidth.

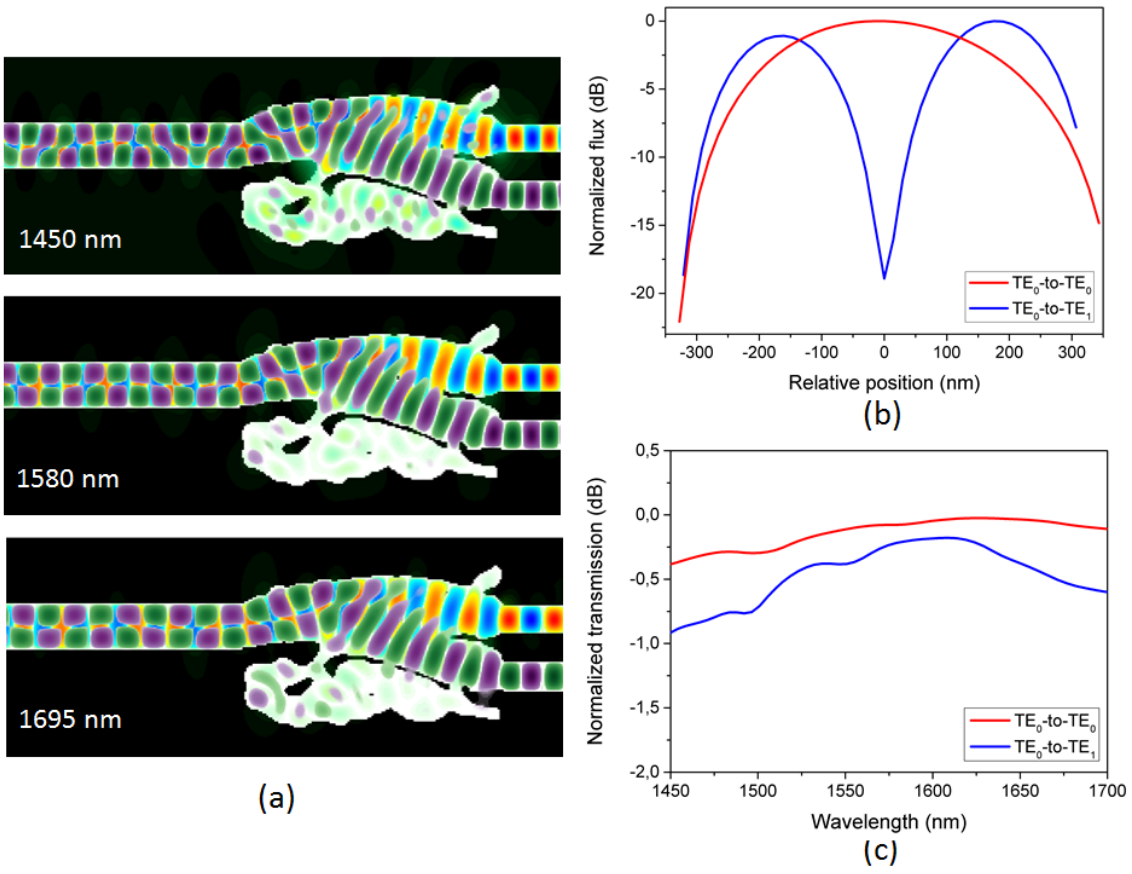


Figure 5.3.12: (a) H_z field simulation through the 3D optimized two mode multiplexer presented in figure 5.3.11 processed at different wavelengths to investigate field distribution in the 'appendix'. (b) Flux of the multiplexed signals at 1580 nm when inducing light in only one of the single mode waveguides respectively. (c) Transmission losses of the 3D 2-mode (de-)multiplexer.

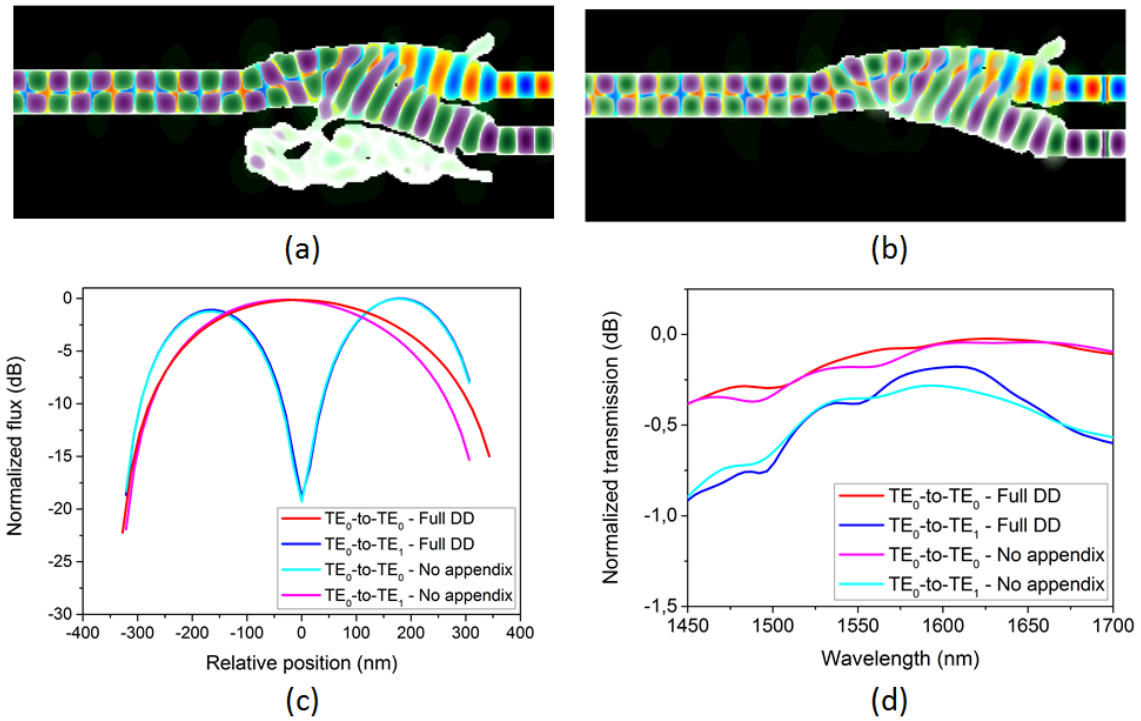


Figure 5.3.13: Numerical comparison of the performance of the 3D topology optimized two mode multiplexer of figure 5.3.11 (a) in its original form and (b) with the 'appendix' manually removed after finished optimization. Both structures overlaid with the simulated H_z field. (c) The simulated mode profiles for light input in either of the narrow waveguides supporting only the TE_0 mode. (d) Simulated transmission spectra for the (de-)multiplexer with and without he 'appendix.

The full structure including the 'appendix' has been fabricated and experimentally characterized. Figure 5.3.14 includes an SEM image of the device along with the collected experimental data. From the mode profiles in figure 5.3.14(b), recorded when light is input in the upper and lower waveguide respectively, it is clear that the device exhibits the desired functionality converting light in one channel while maintaining the TE_0 mode in the other. Figure 5.3.14(c) shows a line-scan across the mode profiles revealing an extinction ratio of at least 17 dB for the center wavelength of 1570 nm. The spectrum reveals a very stable insertion loss underlining the broadband functionality of the device. The loss of a single device is found to be < 1.2 dB while the crosstalk of the mirrored device is less than -12 dB. Comparing this to the results presented for the device in section 5.3.1 the similarity is clear however there has been a quantitative improvement from giving the software a more generic input but with a tuning of the waveguide positions; this in spite of having a slightly smaller design domain. Most likely this is due to the starting features of the previous design guiding the optimization to a less ideal local minimum.

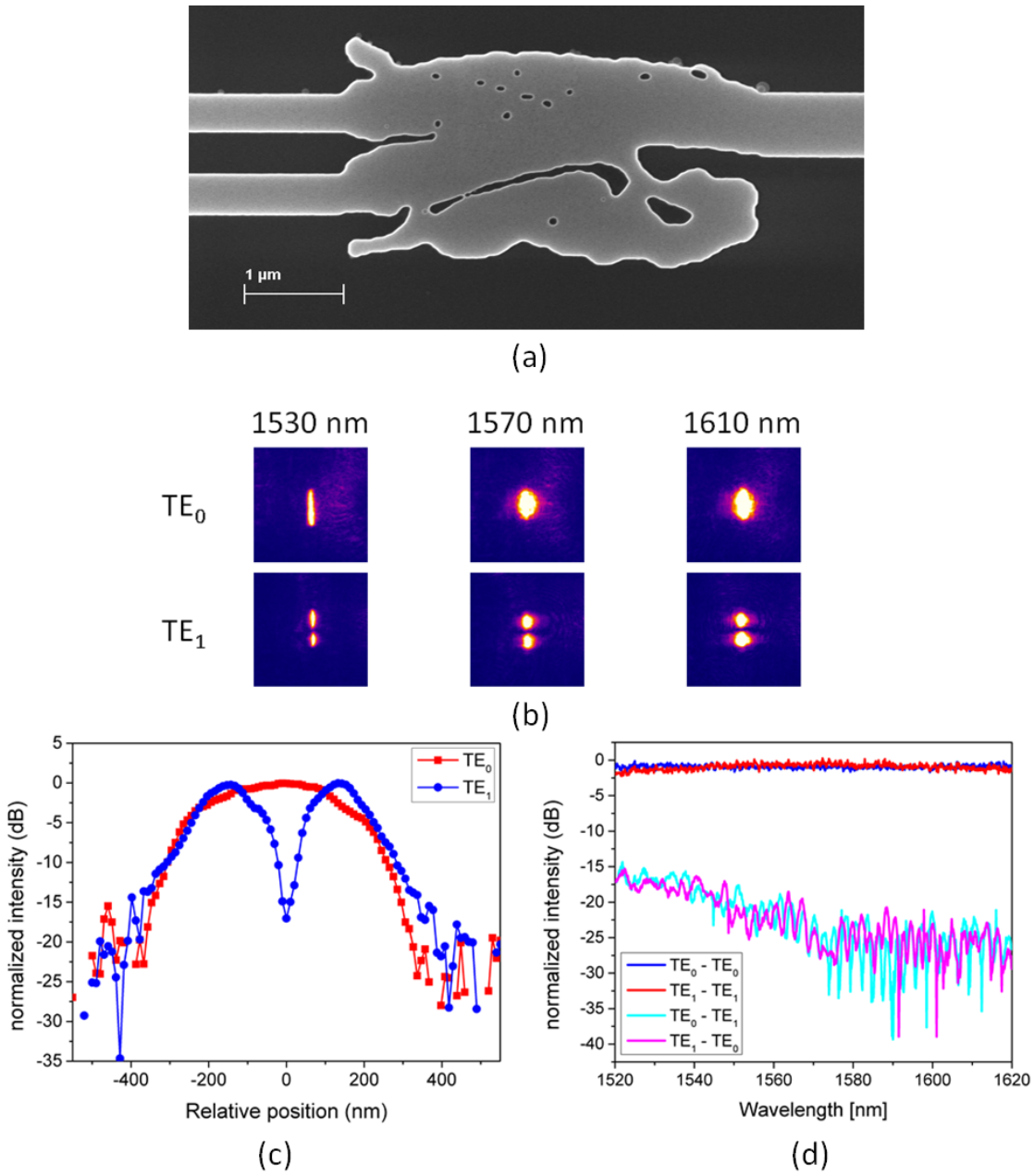


Figure 5.3.14: (a) SEM image of the fabricated 2-mode multiplexer created from a simple box starting point structure. (b) Recorded mode profiles of the multiplexer when light is input in the upper and lower input waveguide, respectively. (c) Line scans across the recorded mode profiles at a wavelength of 1570 *nm. (d) Measured transmission spectra for the mirrored structure with two mode multiplexers, normalized to a straight photonic wire with the same width as the narrow input wires. TE_0 in the legend refers to coupling to the upper waveguide, while TE_1 denotes the fibre coupling light to the lower.*

5.4 3-mode Multiplexer

In this section the goal is to show that one of the benefits of TO is that it can be scaled to include more modes for multiplexing without needing a much larger footprint. A number of the existing schemes can handle many modes, however this is often done by cascading the devices [51, 43] which will lead to a large consumption of chip area. Here a specific 3-mode (de-)multiplexing design is made based on the same principle as described in section 5.3.2. In figure 5.4.1 the starting point structure with a square design domain and three output waveguides is shown. As was the case for the 2-mode MUX system, PSOs are placed in the waveguides carrying only the TE_0 mode and the structure is optimized for minimizing the resemblance to a very low-amplitude pulse shape. A noticeable difference for the wide waveguide, carrying the multiplexed signal, is that it now includes a leading taper. This is because Phazor does not support a mode exciter, this leads to difficulties in introducing the fundamental mode in the wide waveguide. The TE_1 and TE_2 modes are excited by two and three Gaussian sources in the wide waveguide, respectively. Meanwhile the TE_0 mode is excited as a single Gaussian source in the narrow waveguide, with dimensions identical to those of the narrow output waveguides, which is then tapered to the wider region. This causes a need for some further tuning of the time delays for the arrival of the pulses at the objectives but does otherwise cause no further complication. This leading taper is only a workaround needed for the optimization and is not included in the final design, and for this reason simulations confirming the functionality of the device are also made without the taper. The displacement of input and output waveguides are the same as for the 2MUX case with the addition of the longest displacement of 1.472 μm between the input waveguide and the one carrying the TE_2 signal.

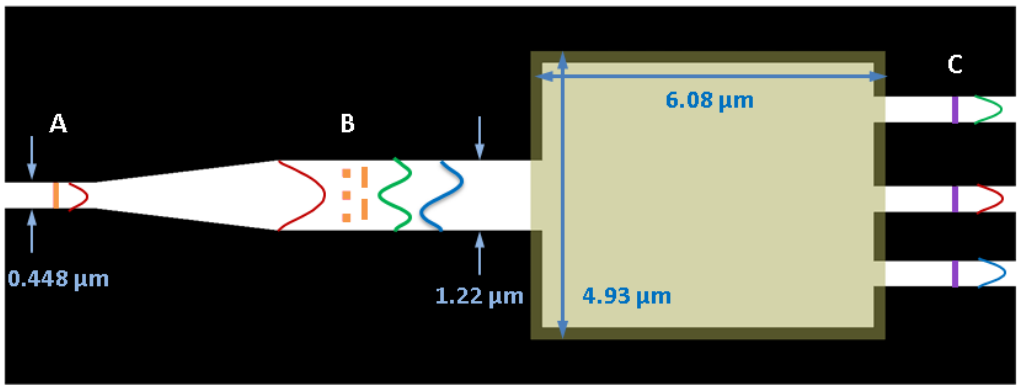


Figure 5.4.1: Starting point structure for the 3-mode (de-)multiplexer operating on TE_0, TE_1 and TE_2 modes. The orange bar in the left narrow region indicates the excitation position for the TE_0 source while the five dots in the wide part of the waveguide represents the excitations sources for the TE_1 and TE_2 modes; they are separated for clarity but both sources are excited at position B. The objectives of the optimization are placed at the purple bars at position C. The resolution for the optimization is 32 nm/pix and the optimization takes place in the square design domain (yellow).

The optimized design is shown in figure 5.4.2 along with field simulations of the three modes; they have been split up for easier qualitative evaluation. The functionality is observed from the field simulations, however it is also apparent that a better starting point, leading to a smaller footprint may be chosen. Especially the lower left region of the structures carries very little of the field. in figure 5.4.3 a structure with the less important regions removed is shown along with simulated transmission spectra of both designs.

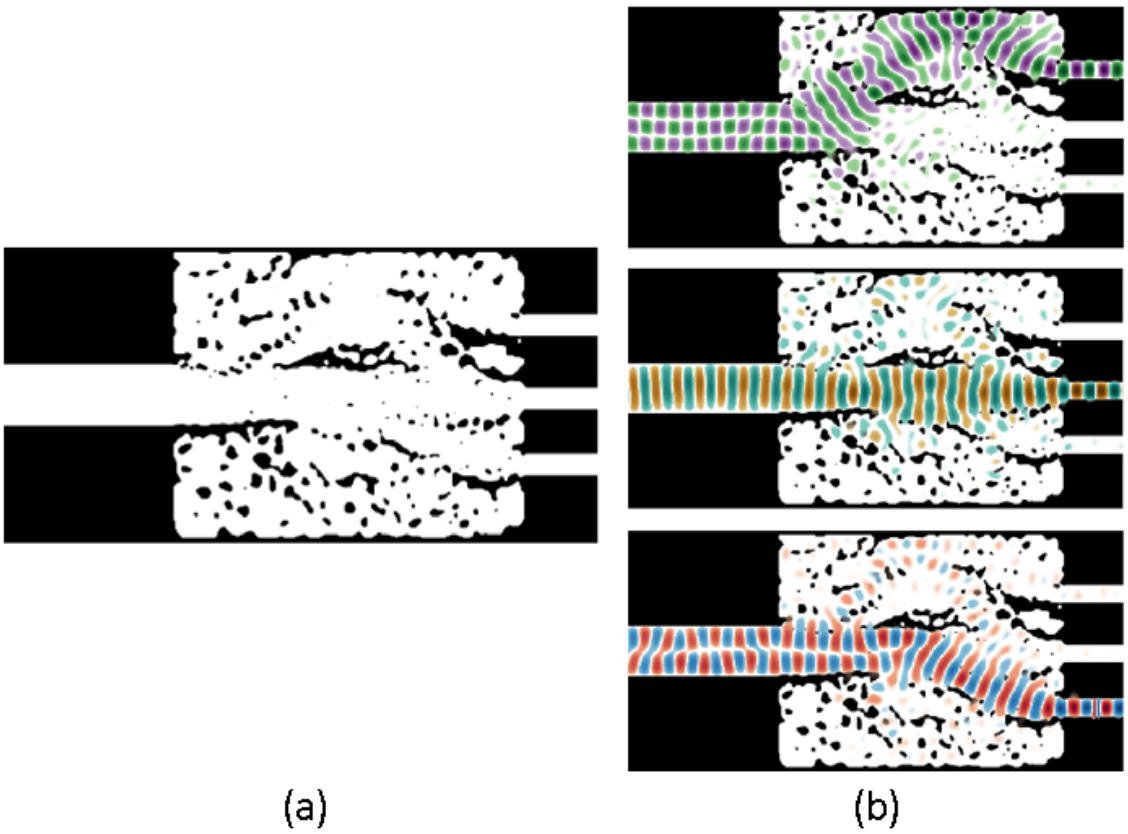


Figure 5.4.2: (a) The topology optimized design of a 3-mode (de-)multiplexer. (b) The simulated H_z field through the structure shown for the three signals excited as fundamental modes separately. The leading taper included in figure 5.4.1 was removed before the simulations were commenced to ensure that it did not interfere with the final device.

The simulated spectra, presented in figure 5.4.3 reveal high functionality for the TE_0 mode, slightly lowered throughput and bandwidth for the TE_1 mode and finally a much lowered bandwidth for the TE_2 mode although the losses near the central bandwidth are comparable to the others. It is likely that this is the case because this design is based on the previous work for the 2MUX device, thus the optimization parameters are more suited for those modes. Further tuning of the parameters might thus lead to a more even distribution of mode performances. Cutting away the side structures is a very simplified method of making the footprint of the device smaller, hence it is no surprise that it increases the insertion losses. It may be noted that there are regions in which the TE_2 mode exhibits higher functionality for the cut-off structure. Presumably this is due to the removal of reflections that cause resonance leading to the high central plateau but also neighbouring performance dips due to interference. Regardless the functionality of the design is maintained when removing the side structures and this indicates that further optimization could lead to a smaller footprint while maintaining low losses. The goal here was to show that scaling the multiplexer to include more modes is possible so no additional steps of optimization are included before fabrication and experimental characterization of the device.

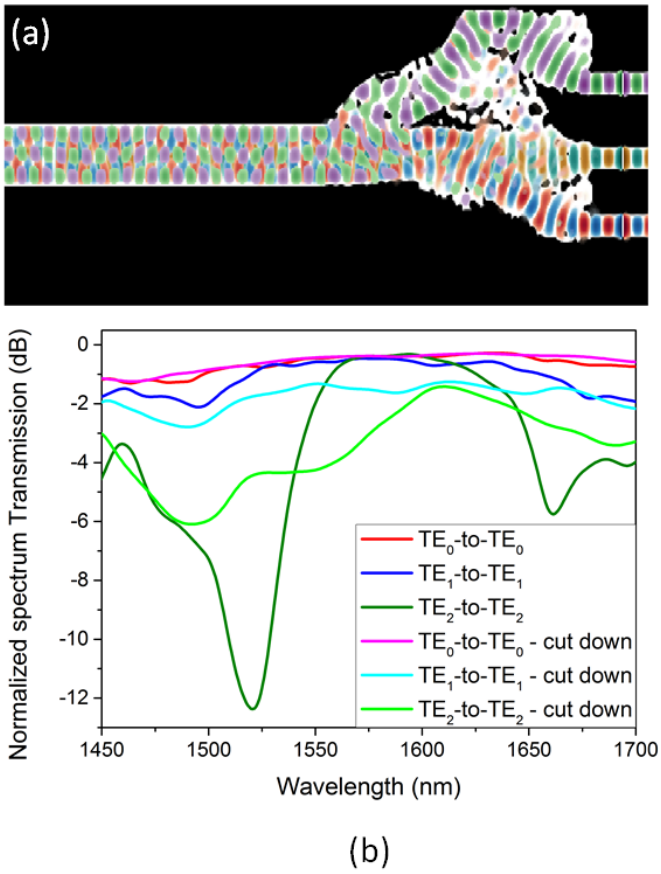


Figure 5.4.3: (a) The topology optimized design of a 3-mode (de-)multiplexer with the parts of the structure where least field resides manually cut away. (b) Normalized simulated transmission spectra of the devices presented in (a) and the one from figure 5.4.2. Spectra have been calculated one at a time when light is input in the narrow TE_0 waveguides.

An SEM image of the fabricated structure is presented in figure 5.4.4 along with the measured spectra of the mirrored structures. For clarity the transmission spectra are shown in three separate plots. Each plot has the light input in the same TE_0 mode waveguide and output at all three possibilities, thus showing the insertion loss and cross-talk of the channel. As expected from the simulation the loss of the TE_0 mode is no more than 0.7 dB for a single device, but the losses of TE_1 and TE_2 are significantly higher around ≈ 2.3 dB. The bandwidth of the TE_2 mode is smaller and is cut off below 1550 nm. Crosstalk varies for each channel but for the mirrored structures it does not exceed 9 dB anywhere within the functional bandwidth of the TE_2 mode.

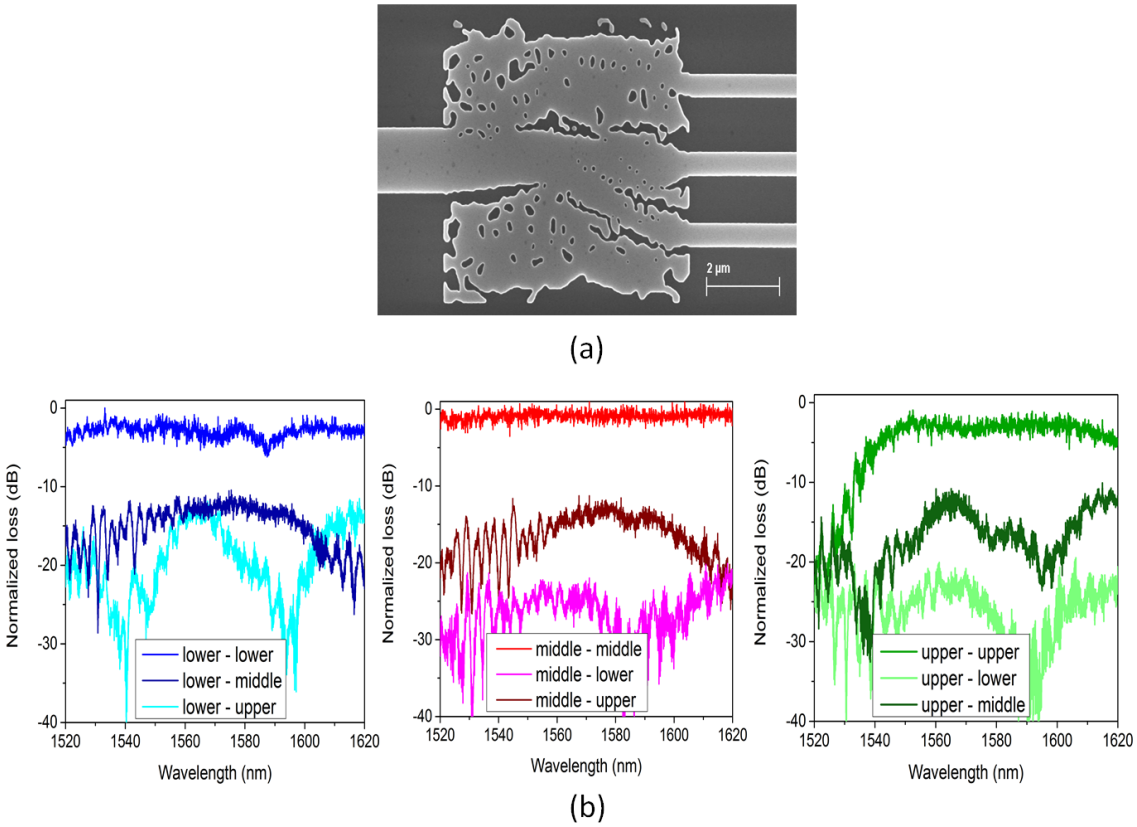


Figure 5.4.4: (a) SEM image of the fabricated topology optimized 3-mode (de-)multiplexer device. (b) Normalized measured transmission spectra of the device shown in three separate plots, each corresponding to input in one of the three single mode waveguides and output after traversing two mirrored structures. Input in the lower waveguide converts to the TE_1 mode, in the middle waveguide maintains the TE_0 mode and inputting in the upper waveguide lead to conversion to and from the TE_2 mode.

Figure 5.4.5 shows the imaged mode profiles from the fabricated device at three different wavelengths. Functionality is present for all modes even at the lower wavelengths where losses are large for the TE_2 mode. The extinction ratio can unfortunately not be determined reliably for this device as the resolution of the IR camera becomes even more limiting for the closely spaced modes. The mode profiles are shown for the TE_1 and TE_2 modes for the same three wavelengths. By lowering the input power for the highest wavelength where the performance of the TE_2 mode is expected to be the best the peak values are lowered but it is possible to get a sense of how good the extinction ratio of the TE_2 mode in fact is.

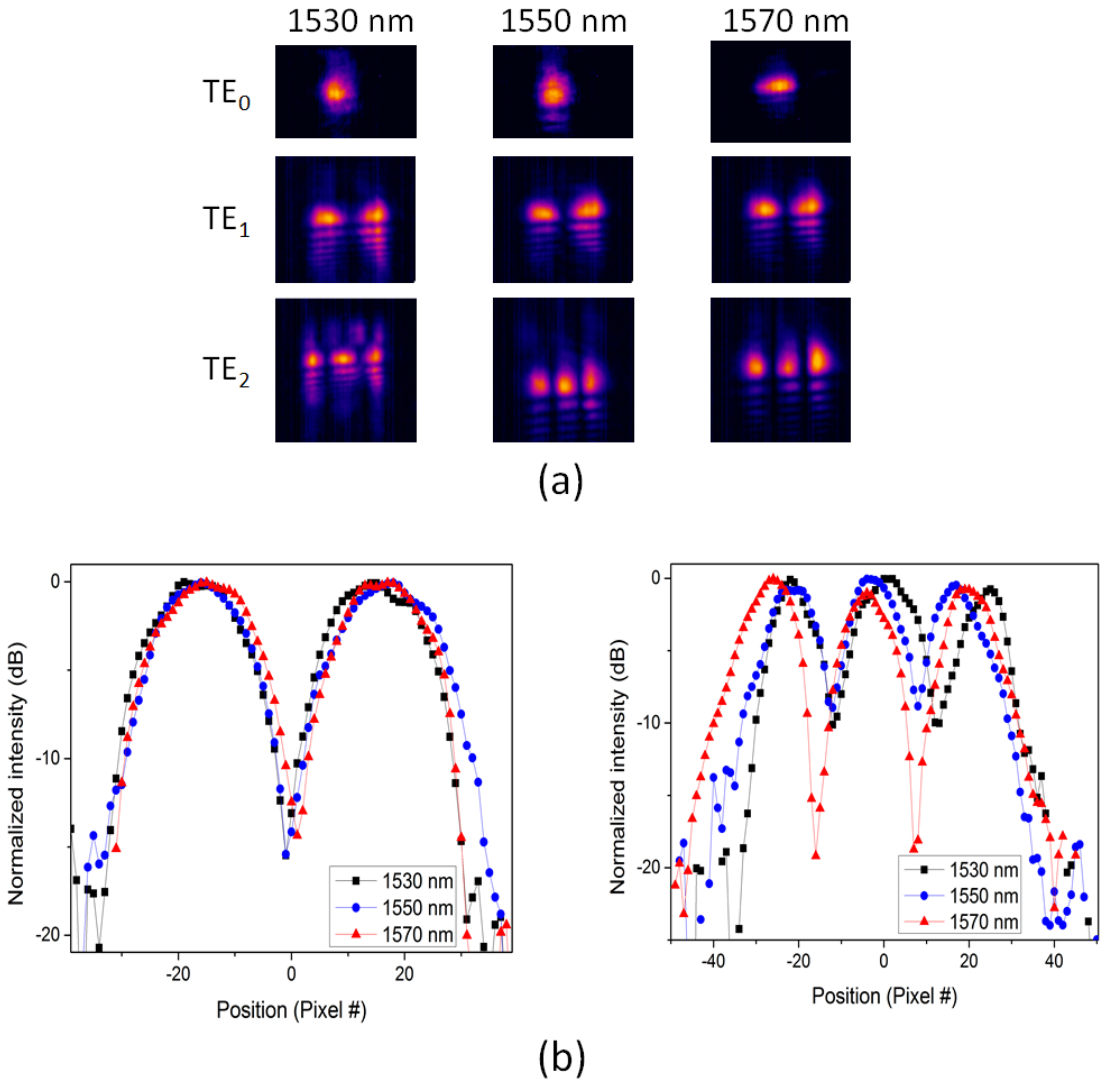


Figure 5.4.5: (a) Recorded mode distribution for all of the channels of the 3-mode multiplexer at three different wavelength. (b) The recorded mode profiles for the same three wavelengths shown for TE_1 (left) and TE_2 (right) modes. The image resolution of the IR camera becomes even more of a limiting factor when the number of modes is increased.

Although it is believed that the functionality could be increased and the footprint decreased if further optimization of the device is commenced, it has now been shown that it is possible to obtain 3-mode (de-)multiplexing functionality using TO without even doubling the side-lengths of the 2-mode multiplexing device.

5.5 Design Robustness

One of the major concerns of the TO method is the robustness and feature sizes of the designs. No difficulties were encountered during this project, as EBL allows for good reproducibility and very small feature sizes. However, these are challenges that must be taken into consideration when investigating the possibility of using TO for designs meant for large scale integration where other fabrication methods must be applicable.

Robustness of a design concerns many different variables such as tolerance to fabrication issues, size changes, and environment variations as well as the general reproducibility. There is no clear definition and hence no direct way of testing for an overall robustness. It is however an important subject and although not the main focus of this project a small scale analysis is included here for completeness. It is relevant to all designs presented in this thesis, however the most thorough analysis, including experimental results, not only simulations, was made with the 2MUX designs, beginning with the assembled design from section 5.3.1. Although this analysis will not give a complete picture of the robustness it provides some understanding of what parameters are important for the design robustness.

Among the approaches taken to investigating the robustness of a design are: Working with different threshold values for the black/white filtering, different resolutions and filter sizes, and investigating the importance of side structures and minor features along with a simplistic numerical investigation of the importance of the temperature.

5.5.1 Resolution and filters

The resolution of the device dictates the possible feature sizes, which can to some extent also be controlled using filters and will be investigated in further detail in chapter 6. Here the performance of a s-mode multiplexer device is investigated for the resolutions of 16 nm/pix and 32 nm/pix, respectively. The structures are shown in figure 5.5.1 along with the measured insertion losses.

As the pixel and filter sizes are determined before the TO, the resulting designs are different even if they carry similarities. For fabrication, larger structures are favourable and it is hypothesized that this is the cause of the losses for the coarser structure being lower. The smaller features present in the high resolution device are, however, still relevant when produced properly and they do significantly lowers the crosstalk of the TE_0-to-TE_1 channel. In the following experimental and theoretical analysis it will be investigated if the increased size generally leads to a better robustness.

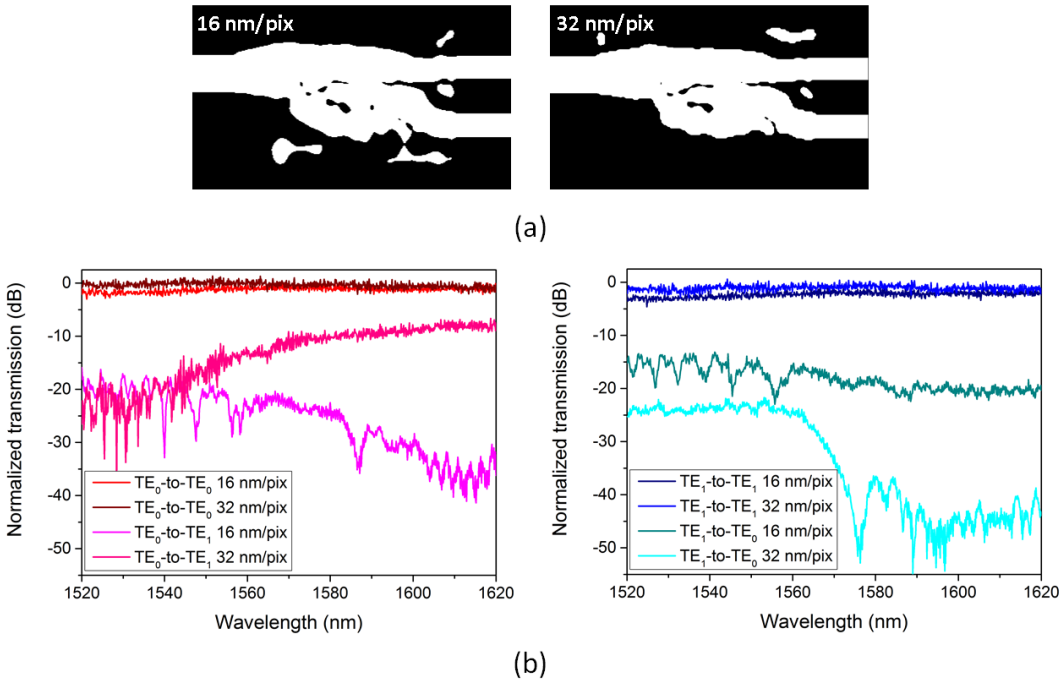


Figure 5.5.1: (a) The topology optimized assembled 2-mode multiplexer with a resolutions of 16 nm/pix and 32 nm/pix, respectively. (b) Recorded transmission spectra for devices fabricated based on the two shown designs (left) when inputting light in the upper waveguide and (right) when inputting light in the lower waveguide; including both intended signal and crosstalk.

Threshold values

Throughout the TO the designs are guided towards binary silicon and air structures, yet there are always some amount of greyscale areas left. When preparing the designs for fabrication a final filtering takes place making the structure completely binary. This is done by defining a threshold value of how dark an area needs to be, ranging between the values of 0 and 1, to be considered black (1) and thus be considered air in the final design. Figure 5.5.2 shows the structure with a resolution of 16 nm/pix for three different values of the threshold along with the experimental transmission losses of these devices when measured in the mirrored device configuration.

While the highest threshold value leads to the best performance for the TE_0 branch of the device, the TE_1 branch operates better for the lowest threshold. This trade-off might be explained by the holes in the design becoming bigger when the threshold value is larger; in this way the signal from the lower branch is more cut off from the top part of the structure where the TE_0 mode resides. For this particular design an intermediate threshold value of 0.35 seems to lead to the best overall performance. The specific value is however something which should be investigated for whichever device is under evaluation as the amount of grey regions and their significance will vary greatly with the specific TO parameters. A more general conclusion about the robustness can, however be made; from the device illustrations it is clear that the changes are relatively small, yet they lead to quite a large change in the performance of the device. Unfortunately this means that for a device like this, made with a high resolution, small alterations can cause significant complications. However the grey areas resides in the regions where the current structure is most affected by change. This means that the same amount of change if distributed evenly or in other parts of the device might not carry the same impact. Robustness to this type of change is investigated later in this section.

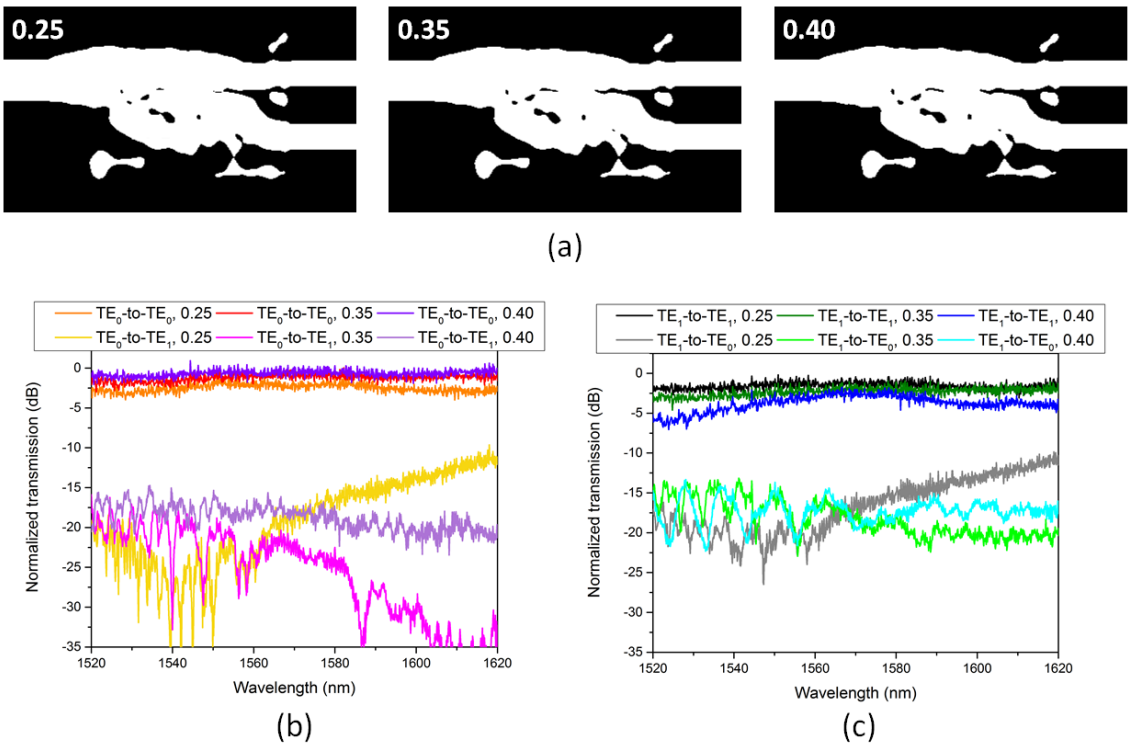


Figure 5.5.2: (a) Illustrations of the combined TO 2MUX, resolution of 16 nm/pix, with three different threshold values of the final filtering. (b,c) Measured insertion losses for the various channel combinations for each of the three design variations. Spectra are divided based on which waveguide the light is coupled to. (b) Light input in the upper waveguide which maintains TE_0 throughout and (b) the light is coupled to the lower waveguide which converts to the TE_1 mode.

Figure 5.5.3 shows the same type of results as presented in figure 5.5.2 but for the device with a resolution of 32 nm/pix. Here the threshold of 0.35 does not reveal the best result, thus underlining, that a theoretical analysis should be undertaken individually for all structures before fabrication. To quantify the robustness the largest relative shift of the insertion losses are summed up in table 5.5.1. The waveguide carrying the TE_1-to-TE_1 signal with a threshold value of 0.25 was damaged and has been disregarded in the analysis. It is found that the robustness with respect to the grey scale threshold value is slightly larger for the lower resolution as would be expected. This is especially the case for the TE_1 mode, however this might be biased due to the reduced dataset.

	16 nm/pix	32 nm/pix
TE_0-to-TE_0	2.24 dB	2.46 dB
TE_1-to-TE_1	4.39 dB	3.50 dB

Table 5.5.1: Approximate shift in transmission loss occurring from changing the grey scale threshold value between 0.25 and 0.4 for the assembled 2MUX device for two different optimization resolutions.

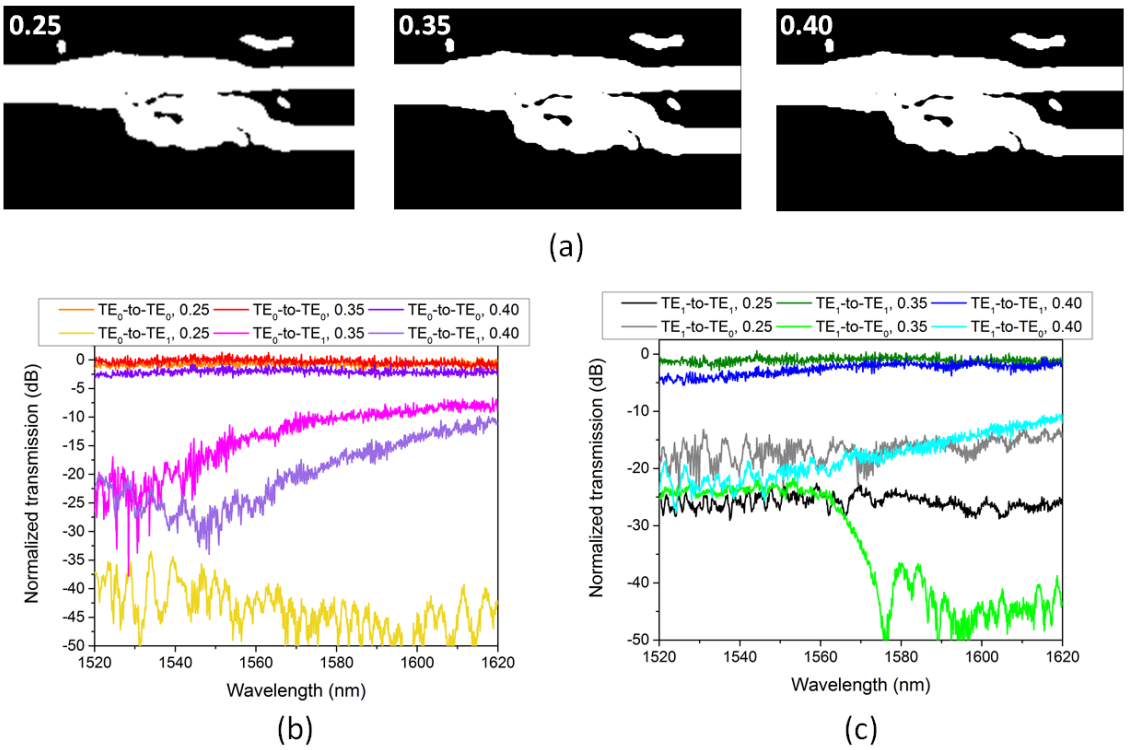


Figure 5.5.3: (a) Illustrations of the combined TO 2MUX, resolution of 32 nm/pix, with three different threshold values of the final filtering. (b,c) Measured insertion losses for the various channel combinations for each of the three design variations. Spectra are divided based on which waveguide the light is coupled to. (b) Light input in the upper waveguide which maintains TE_0 throughout and (b) the light is coupled to the lower waveguide which converts to the TE_1 mode. The waveguide for the TE_1-to-TE_1 signal with a threshold value of 0.25 was damaged, thus severely lowering the overall transmission, it was whoever included for completeness.

5.5.2 Structural changes

During the TO an optimum material distribution is found. Unfortunately there are multiple factors during fabrication that may lead to the structure being changed. Small structures may be difficult to define on the chip and even more so to reliably reproduce. Here the loss of smaller structures as well as uniform size changes, as light be caused by line broadening during etching, is examined through simulations.

Side structures and small features

As has been alluded to in section 5.3.2, not all parts of the device are equally essential. For the islands or appendices outside the main body of the design, it might be apparent that they are not of great importance, but here the influence of the inside small holes is also considered. Figure 5.5.4 and figure 5.5.5 show the performance of the 2MUX assembled design when smaller features are changed. Firstly the side structures outside the functional region of the device have been removed, then the smallest holes within the structures are filled in and finally also medium sized holes are blotted out. Below in figure 5.5.4(d) and (e) the 3D FDTD simulated transmission spectra are shown for the TE_0 and TE_1 modes of the devices, respectively.

It becomes apparent that all features do play a part in minimizing the loss of the multiplexer although to different extents. The smallest features within the multiplexing region play a very small part, which is fortunate as these are the

ones most likely to be lost during fabrication. Although very little field is present in the appendices they do in fact influence the loss of the structure too. This is characteristic of the TO. The method will distribute all material within the DD to minimize the loss. Hence it is likely that a better result will be obtained from a smaller DD in the outset than from removing seemingly irrelevant parts of the device subsequently.

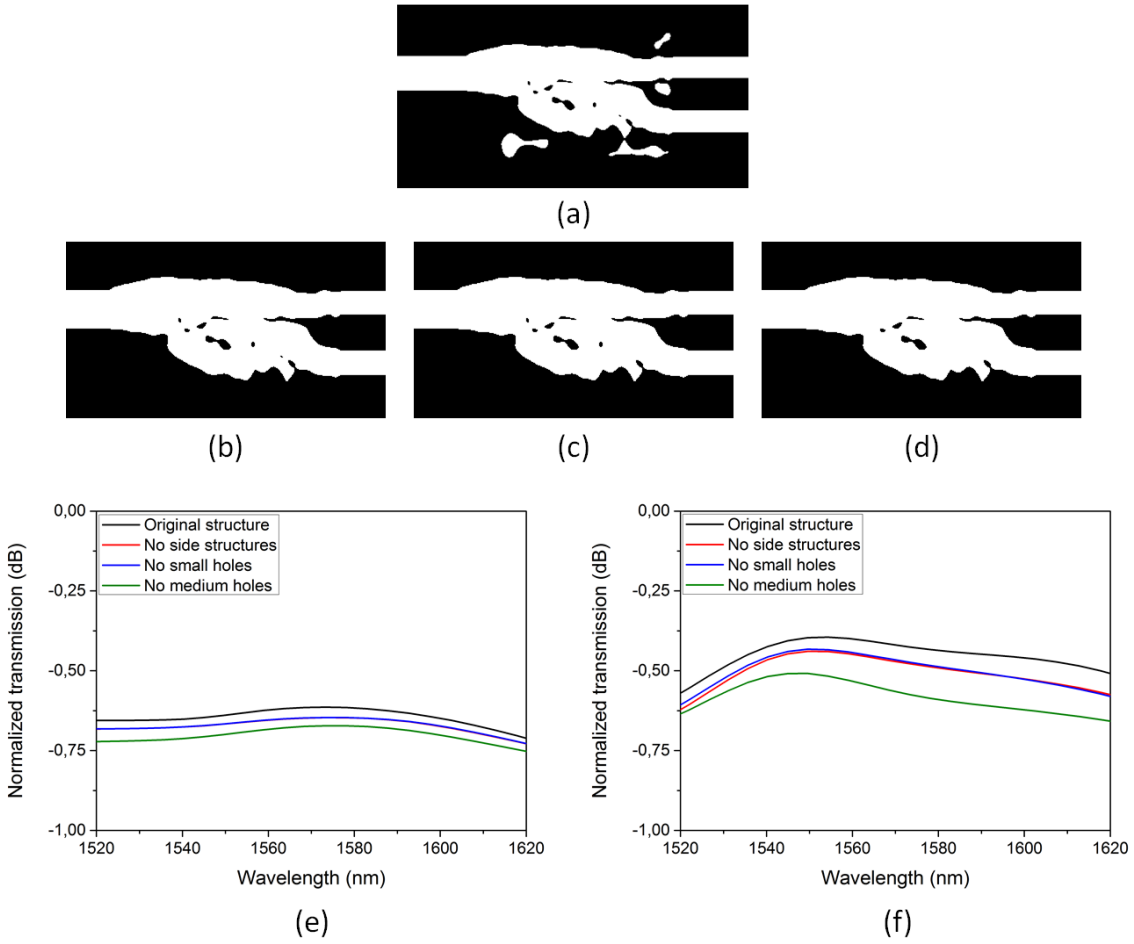


Figure 5.5.4: Illustrations of the combined TO 2MUX, resolution of 16 *nm/pix. (a) The original structure, in the following images altered to exclude (b) side structures outside the main body of the multiplexer, (c) removing also the smallest features inside the structure and in (d) removing also larger features inside the multiplexer region. The 3D FDTD calculated spectra of the four variations of the design for (e) the* TE_0 *mode and (f) the* TE_1 *mode.*

Comparing the two resolutions under investigation does not reveal a clear correlation. The biggest qualitative change in the spectrum is observed for the 32 nm/pix case. The quantitative difference for the TE_1 mode is very similar for the two resolutions, however the loss is larger for the lower resolution when concerning the TE_0 mode. Although one might expect the lower resolution with bigger features to be more robust, it is to be expected that removing features from it will have a larger impact. When the features are larger in general, removing even just the smallest features does still constitute a larger change area wise, and it is thus natural that it has a larger impact.

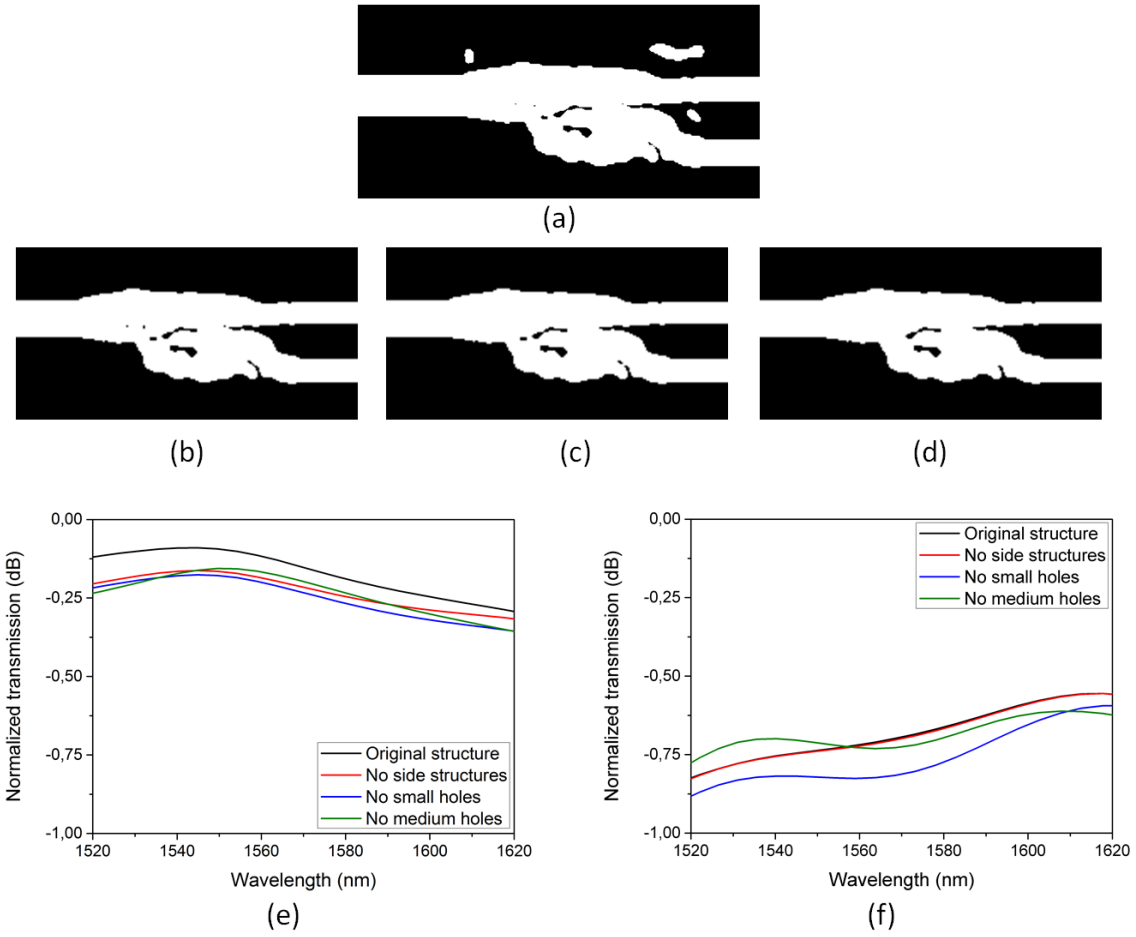


Figure 5.5.5: Illustrations of the combined TO 2MUX, resolution of 32 *nm/pix. (a) The original design, in the following images altered to exclude (b) side structures outside the main body of the multiplexer, (c) removing also the smallest features inside the structure and in (d) removing also larger features inside the multiplexer region. The 3D FDTD calculated spectra of the four variations of the design for (e) the* TE_0 *mode and (f) the* TE_1 *mode.*

Uniform size changes

When changing the threshold-values of the structure the features of the design are altered to various degrees in the areas where the device sensitivity is highest as it is not yet fully converged. Although one can never know what types of fabrication errors may occur, it is likely that uniform size changes will take place. A uniform increase of sizes is in fact expected during the etching procedure and the structure features are shrunk to account for this. In this section the effect of a uniform size change is investigated, both for shrinking and expanding the full structure.

Figure 5.5.6 presents the spectra for the 2-mode multiplexer as it is shrunk and grown, respectively. As would be expected, the changes affect the TE_1 mode the most. Shrinking the structure, which causes the holes inside the device to grow bigger, has a stronger impact than increasing it uniformly. In both cases changes of ± 5 nm lead to only a relatively little increase in the loss. Changing the size further does make the effect more severe and doubling the change to ± 10 nm leads to a doubling of the loss as well. Unfortunately, this does not indicate a strong robustness of the structure. For EBL these are acceptable limits, but it does show that more work is needed before other fabrication procedures would be applicable. Furthermore this investigation reveals the importance of correcting for the line-

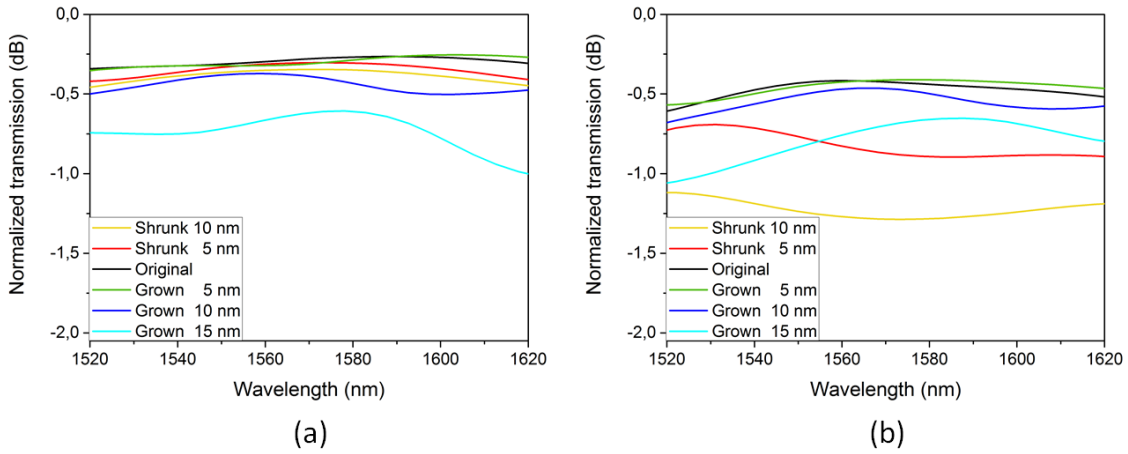


Figure 5.5.6: Spectra for the 2-mode multiplexer with a 16 nm/pix resolution as the structure size is changed uniformly. (a) For the TE_0 mode carrying channel. (b) For the channel converting from TE_0 to TE_1.

broadening when preparing the etch mask.

5.5.3 Temperature variations

Although the temperature is kept stable in the lab conditions under which the devices are tested, temperatures would rise in an integrated electro-optic system which such designs are envisioned to be part of. For elevated temperatures starting from around 25 °C the temperature dependence of the refractive index is linear[7]:

$$n(\lambda, T) = n_0(\lambda) + a_n(\lambda)T, \tag{27}$$

where $n_0 = 3.4176$ is chosen to match Phazor's standard value of the refractive index at room temperature. The coefficient of the linear term is a function of the wavelength following the fifth-order polynomial:

$$\begin{aligned} a_n = 2.46796 \times 10^{-2} - 1.48108 \times 10^{-4}\lambda + 3.51075 \times 10^{-7}\lambda^2 \\ -3.99478 \times 10^{-10}\lambda^3 + 2.15422 \times 10^{-15}\lambda^4 - .35668 \times 10^{-17}\lambda^5. \end{aligned} \tag{28}$$

It is estimated that realistic operating temperatures should be between room temperature, 20 °C, and 100 °C[3]. Figure 5.5.7 shows the relation between the refractive index and the temperature. Simulations have been made for both of the resolutions under investigation and in figure 5.5.8 the insertion loss at the center frequency is shown as a function of the temperature. Along with this the unnormalized spectra are shown to illustrate the effects of the temperature change.

Unfortunately it is apparent that the designs do not maintain their low losses throughout the entirety of the range and especially the TE_1 mode efficiency deteriorates.

Minor variations of a few degrees have a relatively small effect on the performance, as is clear from figure 5.5.8(d). The design is thus perfectly stable within the range of operation during this project. The dependence does however indicate that it is important to optimize the structure for the specific operation temperature and keeping the conditions stable close to that.

[3]Based on current operating temperatures of CPUs

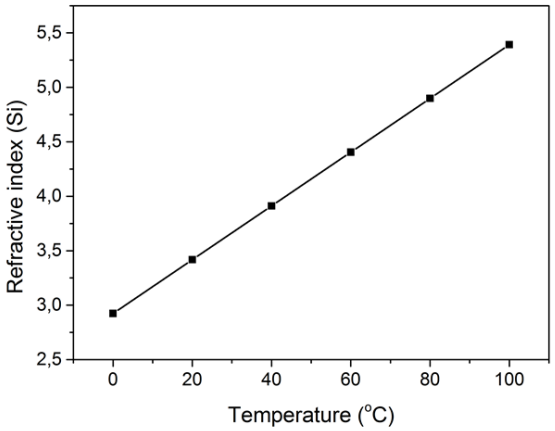


Figure 5.5.7: The refractive index of silicon as a function of the temperature, based on equation (27) *and* (28) *originating from [7].*

It should be noted that this analysis does not account for the thermal expansion of the silicon, nor that there is also a change of the refractive index of the silica layer. As such it is only a first step in a temperature-dependence analysis. This subject is not taken further within this project, but it is of importance for future designs which are to be integrated with electronics.

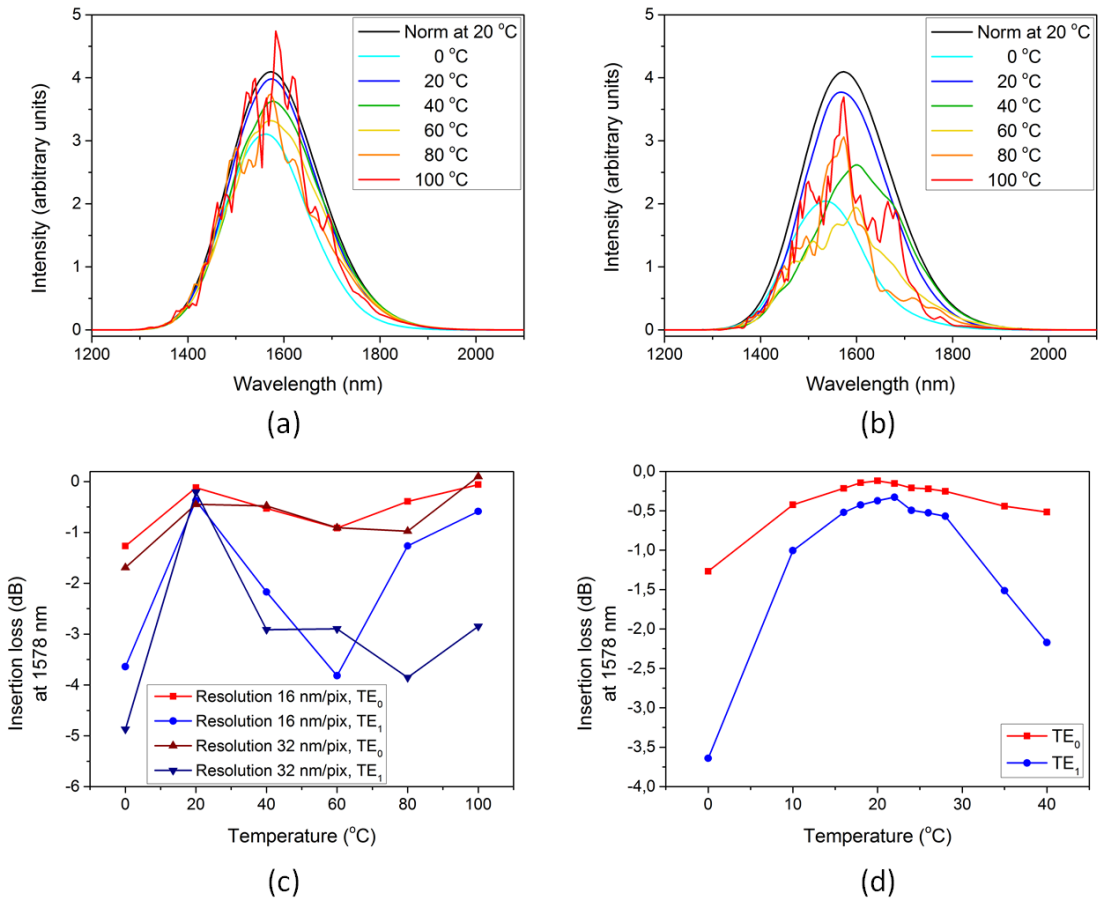


Figure 5.5.8: Simulated unnormalized spectra of the 2MUX device obtained at various temperature for the high resolution design (a) for the TE_0 mode and (b) for the TE_1 mode. (c) The insertion loss of the combined 2MUX device at the center wavelength as a function of the temperature shown for both modes of the 16 nm/pix resolution and 32 nm/pix resolution. (d) A finer spaced temperature investigation for the 16 nm/pix around the central temperature of 20 °.

5.6 Summary and outlook

In this chapter the design and characterization of various components for mode division multiplexing has been introduced.

Initially a mode converter in a PhC was introduced before the concept was tested for a slab waveguide. It became apparent that it was not possible to create a compact functional mode converter in a straight PhW, but that introducing a height difference in the waveguides and optimizing the region in between allowed for excellent performance devices. Iterating from the optimized design, it was possible to obtain an even more compact structure with low losses and good conversion, but a slightly lower bandwidth than its larger counterpart. Comparing the PhC mode converter with those realized in slab waveguides it was found that optimizing directly for the PhW the device region could be shrunk from 22.68 μm^2 to 4.76 μm^2 while lowering the insertion loss from $\sim$ 2 dB to $\sim$ 1 dB in a $\sim$ 44 nm bandwidth. This underlines the point that giving the TO more freedom can lead to increased functionality.

Finally it was shown, through simulations, that knowing the basics for the TE_0-to-TE_1 mode converter, a more complex device working on a higher order mode was easily designed.

Subsequently the more complicated device, namely a mode (de-)multiplexer, was investigated. Numerous iterations were made and a good starting point found by combining the designs most promising for the TE_0 and the TE_1 mode respectively. From this a well-functioning design was found and was experimentally verified to exhibit low loss and crosstalk. This design was improved by using a simpler starting point structure for the optimization whereby the TO algorithm was given more freedom and converged to a better local optimum. This design underwent substantial fine-tuning which led to experimental insertion losses of $< 1\,\mathrm{dB}$ and crosstalk below $-12\,\mathrm{dB}$ for a $4.60\,\mu\mathrm{m} \times 4.22\,\mu\mathrm{m}$ device. Just as for the mode converter it became apparent that some regions of the design were less essential to the functionality than others and a future investigation might include optimizing on a cropped starting point.

The strength of the TO process was highlighted by the creation of a 3-mode (de-)multiplexer which was based on the same principles as the one working on 2 modes. Although the design was not as finely tuned as the previous one, leading to higher losses, it demonstrates the versatility of the tool. It is not necessary to cascade the structure and the designer does not need to consider an entirely new design to obtain more complex functionality.

In the work with both the mode converters and the mode multiplexers it has become clear that a good starting point structure is key for a satisfactory device. Topology optimization is an iterative process - and so is setting up the optimizations. By learning from the early optimization attempts, even if they are suboptimal, it is possible to guide the optimization to a favoured optimum. Both 2D and 3D optimization was investigated in this chapter in the hopes that the faster 2D optimization could help create a good offset for subsequent 3D optimization. Unfortunately, even employing the effective index approximation, 3D optimization is a necessity for good correspondence between simulations and experimental results. Optimizing in 2D may still reveal interesting information on the performance of the devices, however fine tuning in 2D is not profitable as the optimizations parameters are not transferable. The optimization method will always do its best to fulfil the objective, and thus the decision of whether it is footprint, insertion loss, crosstalk, large features that are easy to fabricate or something else entirely which is most important for the present design, and guiding the optimization to favour that, is the important task of the designer.

The strength of TO is that it can deliver record small devices as presented here. The challenge of this, as has been made evident in the last section of this chapter, is that small devices require small features and this lessens the robustness to fabrication errors. In general the TO will lead to every pixel of the final design being carefully chosen. This does not mean that functionality is lost, when the structure is lightly altered during fabrication, but it does mean that almost any change will lead to deterioration of the performance. Because of this it will be an important next step for TO to implement the search for robust structures into the optimization algorithm. An initial step in that direction is to use filtering to ensure larger features, which will be easier to fabricate with high fidelity. These filtering mechanisms is one of the key points of the next chapter.

CHAPTER 6

SIZE-PERFORMANCE-CORRELATION IN WAVELENGTH MULTIPLEXERS

6.1 Introduction

Whereas mode multiplexing, dealt with in the previous chapter, is an emerging technology with the potential of increasing data capacity, wavelength multiplexing is already a well-established method integrated in fibre networks [3][54]. It is thus also a technology of interest in the field of on-chip devices and multiple schemes have been proposed including ring resonators[55, 56], echelle gratings [57], cascaded Mach-Zehnder [58] devices and arrayed waveguide gratings [59, 60]. They are all based on interference between parts of the input light, which can be altered by different propagation lengths. As with most PICs footprint and robustness to fabrication errors are challenges of the various designs.

In the **ring resonators**, illustrated in figure 6.1.1(a), the light can be coupled from a bus waveguide to rings being resonant only with a specific part of the spectrum. The coupling and resonance can be tuned by distance from bus waveguide to the ring, the radius and the width of the ring waveguide which alters the effective refractive index. Heaters can be incorporated with this design to control the dimensions and thus tune the structures after fabrication [55]; the spectral response can also be altered by the number of rings present[56].

The **echelle gratings**, sketched in figure 6.1.1(b) has a single input waveguide and numerous output waveguides ending in a free space region. The light freely diverges until it impinges on a number of reflectors carefully placed as a diffraction grating. The reflectors are adjusted so that the light is directed into a specific diffraction order. As the reflection angle of higher order refracted beams is wavelength-dependent, light can be focused so that different wavelengths will select different output waveguides positioned next to each other[57].

The **Mach-Zehnder type** wavelength multiplexer creates a lattice filter by including a separate delay line consisting of half-circles spaced with a small straight sections in between, allowing for delay adjustment. The main benefit of this structure is that it can increase the free spectral range (FSR) but in addition the bandwidth can also be controlled through the number of periods [58]. The concept is illustrated in figure 6.1.1(c).

The concept of **arrayed waveguide gratings** is illustrated in figure 6.1.1(d). They consist of two slab waveguides and a number of different length arrayed wave-

guides with a constant path length difference between adjacent waveguides. In the waveguide array the delay causes different phase shifts and the light will interfere in the second slab waveguide leading to the (de)multiplexing [59].

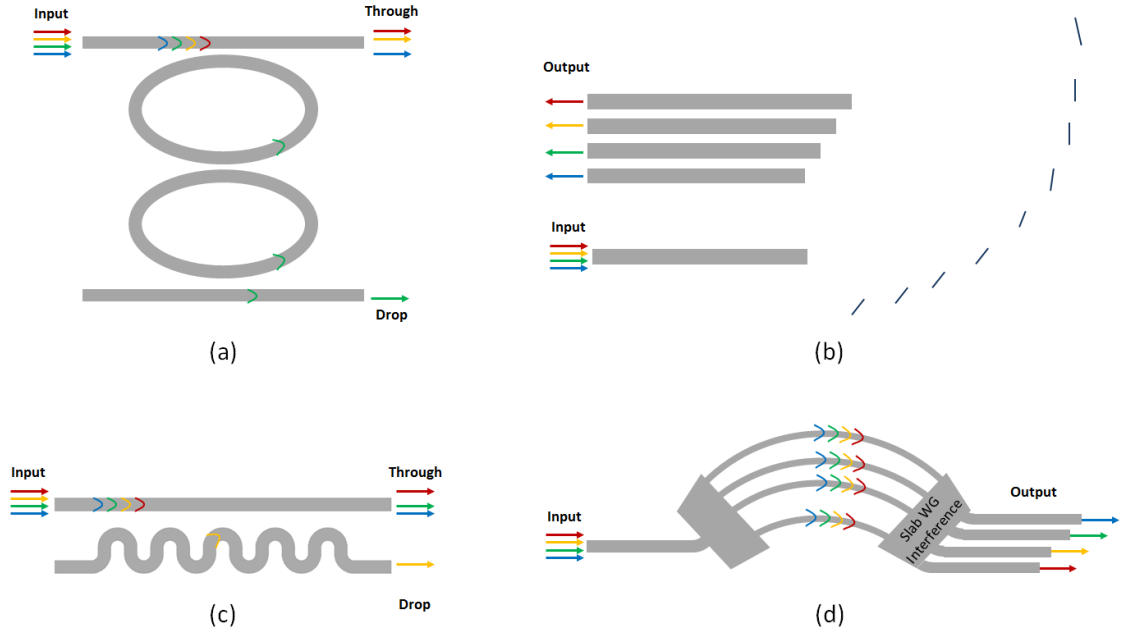


Figure 6.1.1: (a) Schematic structure of a ring-resonator based WDM system. (b) Schematic structure of an echelle grating based WDM system. (c) Lattice filter based WDM system. (d) Arrayed waveguide grating based WDM system.

The sizes of these devices range from tens to hundreds of μm^2 and terms like compact, ultra-compact and ultra small show up in the description of many of them. This is not unjustified as these devices have all pushed the limits of the footprint of the structure type. However there is no definitions of what these terms cover. This type of device can also be realized with inverse design methods which has already been accomplished. The first device was presented in 2007 in a PhC waveguide [37] using TO and later redone in a photonic waveguide in 2015 [22] using the objective first method. These methods allow for yet smaller footprints, however it does not illuminate if the limit has been reached. Because of the complex scattering mechanisms of the inverse designs it is not straight forward to express theoretical limits of the size, yet is is of interest to map out how small devices can be realized. The hypothesis of this chapter is that many devices could be reduced in size but at the cost of the performance. The goal is to obtain a better understanding of the trade-off between footprint and performance by making a systematic investigation of the correlation between the footprint, feature sizes, and performance of an on-chip device. This chapter thus deals with designs of a wavelength (de)multiplexer, not for the sake of the device itself, but because it is one of the easier devices to realize with the TO PEO. It is believed that the trends found for this device will carry over to other topology optimized designs.

This chapter presents a study of feature sizes and performance. Firstly the configuration and design set-up will be introduced. Then the effects of footprint will be investigated, followed by an observation on the importance of the starting point structure, before a study of the importance of the feature sizes within the device. Before summarising and concluding upon the findings, a cursory comparison will be made to the results obtained elsewhere.

6.2 Size-performance-correlation

In the investigation of the footprints effect on the performance of an integrated device any limiting factors are first considered. In this section the fabrication constraints to feature sizes have been disregarded, it will however be reconsidered in section 6.3. To determine the smallest possible structure to investigate, the coupling distance of the output waveguides has been taken into account. If the waveguides are close enough that they may couple, more space would be needed on the chip, to lead the waveguides away after conversion, hence this is a fixed parameter determining also the smallest possible height of the device. The results of the simulation are presented in figure 6.2.1. The calculation has been done for two 500 nm waveguides, the actual width may however vary by a few nanometers to accommodate specific resolution needs for the optimization. The minimum distance between the waveguides used will be 700 nm edge-to-edge as this will ensure less than 1 ‰ coupling efficiency even for high wavelengths.

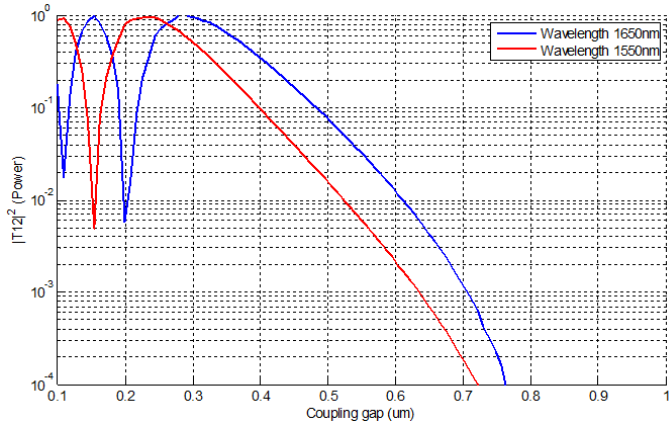


Figure 6.2.1: Result of the simulation of the coupling strength between two 500 nm waveguides. Simulation kindly supplied by Yunhong Ding.

With the waveguide distance determined, the starting point structure can be designed. The base structure for this investigation is a square region of varying dimensions with one input waveguide and two output waveguides symmetrically placed as shown in figure 6.2.2. Side lengths within the range of 1.7 μm to 5.14 μm have been optimized. Several altered structures will be presented later using rectangles rather than squares and varying the waveguide positions.

The final step of preparation is done to ensure that the objectives are weighed evenly. Given that the two wavelength spectra do not carry the same amount of energy the objectives should be weighed accordingly. The spectra are shown in figure 6.2.3, the integrals of the spectra give the values:
1300 nm: $-6.4580867524892 \times 10^{-5}$
1550 nm: $-7.5035711618459 \times 10^{-5}$
As there is less power in the 1300 nm pulse its weight will be 1.162 times greater than that for the 1550 nm objective.

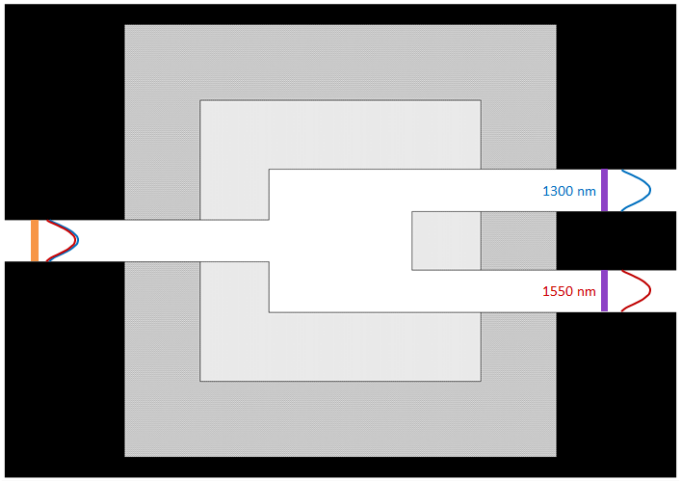


Figure 6.2.2: The starting point structure for the symmetric wavelength multiplexer. Variations shown in overlay to clarify the difference, presenting design domain and structure sizes of $1.7^2\,\mu m^2$ *(light grey structure),* $(3.34\,\mu m)^2$ *(dark grey structure) and* $(5.14\,\mu m)^2$ *(black structure). The resolution of the design is* $20\,nm$ *and waveguide widths all 500 nm.*

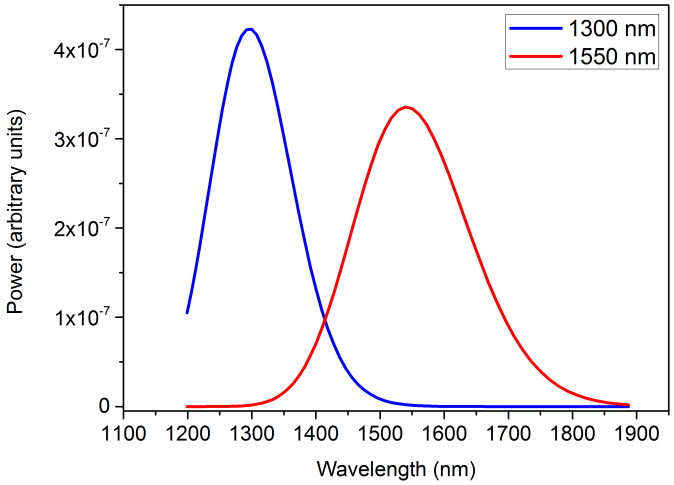


Figure 6.2.3: Raw spectra of the two pulses at $1300\,nm$ *and* $1550\,nm$ *respectively. As the higher wavelength pulse has a higher energy the weights will be scaled correspondingly.*

6.2.1 Square structure

The first step in the investigation of the influence of the size has been to compare the performance of square DD structures with varying side lengths as illustrated in figure 6.2.2. Examples of the optimized designs are shown in figure 6.2.4(a) along with a comparison of their performance in 6.2.4(b). As described in section 2.1.1 the objective function is a measure of how well the optimization has been able to fulfil the objectives. When sources and objectives are identical for various starting point structures, the objective value can be used as a figure of merit when comparing the optimized devices. Here the objective value of the designs have been normalized to that of a straight waveguide with the same sources and objectives. When optimizing a straight PhW for the lowest loss propagation, it will be unaltered and the objective value will be a constant and a numerical description of the best possible design. Figure 6.2.4(b) illustrates the evolution of the objective function after 198 iterations as a function of the size of the square DD.

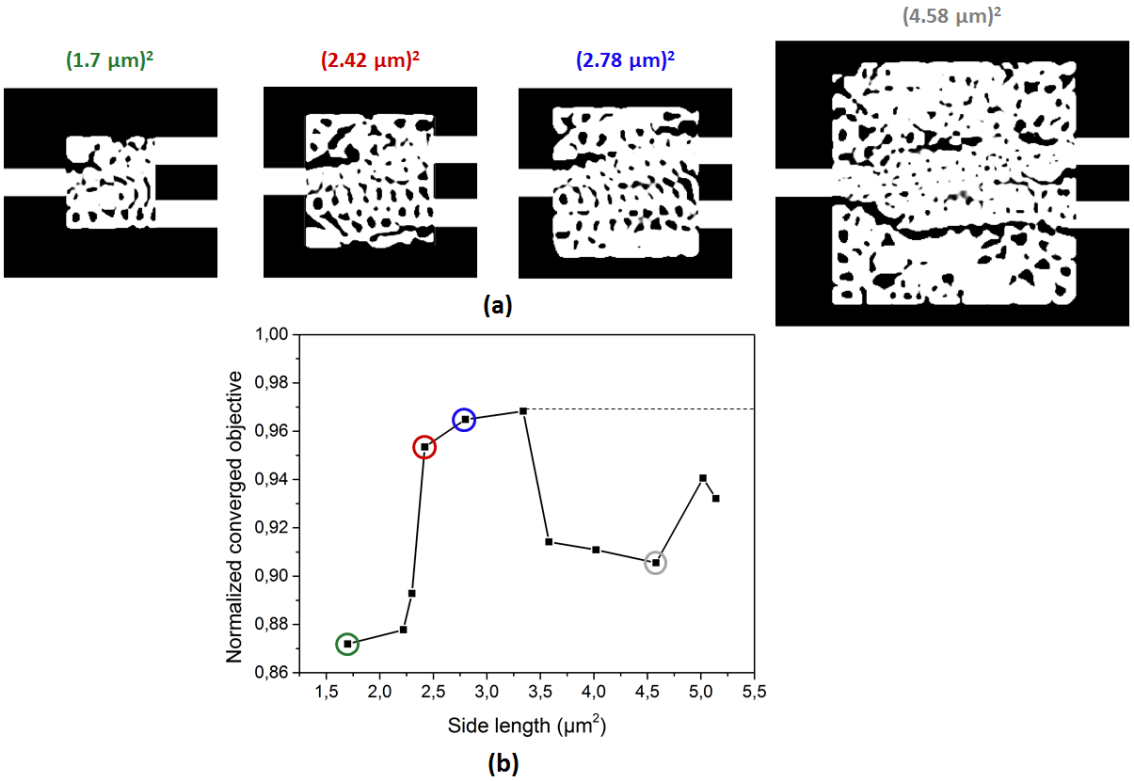


Figure 6.2.4: (a) Topology optimized wavelength multiplexers obtained with the same optimization parameters but varying the size of the square design domain. All depicted in the same scale. (b) Normalized objective function value correlated to the design domain area, the connecting line is just for guiding the eye and the coloured circle for easier identification of the above structures. The dashed line indicates the minimum obtainable best performance in the region.

To verify that the objective function is a good indication of the performance, and thus that the optimization has actually been successful, the normalized spectra of the structures are presented in figure 6.2.5.

From the spectra it can be seen that even the structures exhibiting the poorest objective values do function. With the degradation of the performance most prominently manifested in the extinction ratios, and only to a smaller degree in the insertion loss.

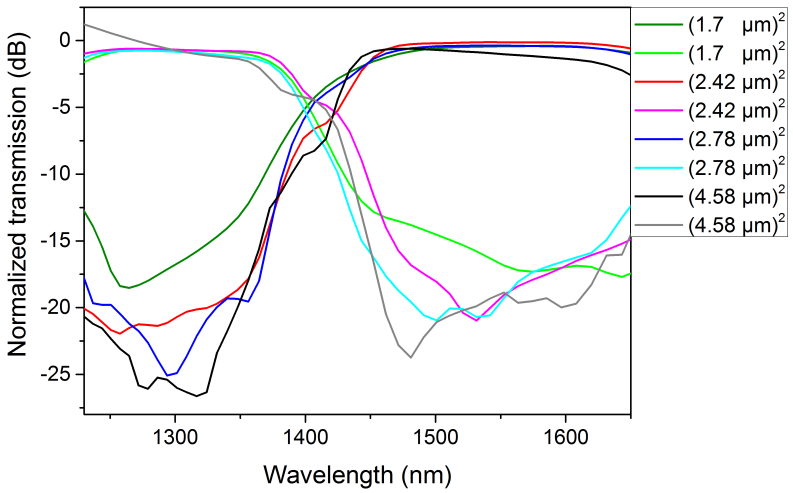


Figure 6.2.5: Spectra of the topology optimized wavelength multiplexers presented in figure 6.2.4(a). All spectra are normalized to the spectrum of the same source propagating through a straight waveguide.

There seems to be a specific size range in which the objective is closest to the optimum and hence the performance best. Within this range a larger size seems preferable, however as discussed in chapter 1 chip area is a valuable resource and this means that it will be necessary to prioritize between lowering the footprint or the transmission loss. That very small structures exhibit poor performance is a result of a physical limit. If the optimized area is too small for the necessary scattering and interference to take place, the performance will be lowered. The lowering of the performance for larger structures is, however, and expression of the math and how the software converges towards a local minima. For any DD there will always exist one solution which is the same as the configuration of any of the smaller DDs. The dashed line on figure 6.2.4(b) indicates this relation. Meanwhile that the TO tends to converge towards a design with a worse objective fulfilment insinuates that the larger area is not much preferable. Hence the high objective value region is still considered an optimum for the trade-off between footprint and performance. It is crucial to keep in mind that due to the nature of the iterative design processes, all of these results are specific local optima and other similar or different designs may exist with very similar functionality. It takes only small changes in the set-up to cause the system to converge to a different optimum and thus this analysis would need to be remade, whenever a device is optimized, to ensure the same settings.

The importance of the starting point for the specific convergence is highlighted in figure 6.2.6. Here the well-functioning $(2.78\,\mu\text{m})^2$ structure from figure 6.2.4 has been inserted as a starting point for the, previously less successful, $(4.58\,\mu\text{m})^2$ DD case. The structure obtained from this TO is clearly differing from the one presented in 6.2.4(a), and the convergence of the objective value is much better as shown in purple in figure 6.2.6. The dashed line added to the figure indicates the presumed best case relation for finding global optima. This makes it evident that larger DDs can give better performance if the starting point is tweaked to nudge it to the correct local optima. It will thus always be necessary to make the choice of which losses and crosstalk can be tolerated to obtain small footprints.

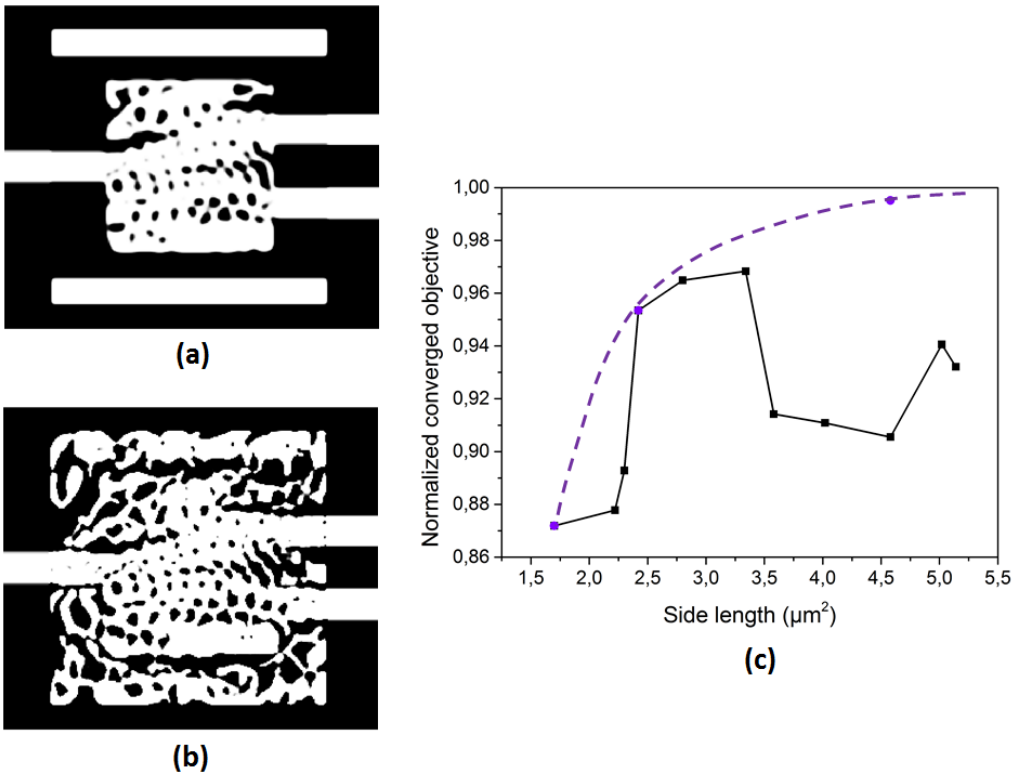


Figure 6.2.6: (a) Starting point structure for the re-optimization of a large $(4.58\,\mu m)^2$ *design domain wavelength multiplexer. Here the central part is an already well-functioning optimized structure with a smaller footprint of* $(2.78\,\mu m)^2$ *(b) The topology optimized structure obtained from the refined starting point shown in (a). (c) Normalized objective function value correlated to the design domain area, same as presented in figure 6.2.4(b), but with the addition of the value obtained from the re-optimized structure presented in (b). The purple dashed line indicates the presumed development of the correlation if designs closer to the global optimum of each device were found.*

6.2.2 Rectangular structure

In the previous section it was found that there are specific footprint ranges which lead to enhanced performance. In this section the possibility of lowering the footprint by decreasing only one of the DD sides will be examined. It may be that either the direction parallel or vertical to the propagation of the light is more important for the scattering design, thus a number of variations have been made to map out any potential correlation. The resulting structures are presented in figure 6.2.7 and the corresponding normalized converged objective function values are shown as a 3D colour map in figure 6.2.8; the sizes have been chosen in the vicinity of $2.78\,\mu m$ which had a promising functionality.

All of the structures have similar characteristics of creating two guided regions to the output branches, with an excess of holes in the region between them and a scattering area on the outside. The larger structures definitely exhibit higher objective values. Across the full size range it appears that having a wider region leads to a better result. This is intuitive, even though there is scattering inside the device, there are no structures directed at disrupting the propagation and altering the direction, so the vertical direction becomes secondary. In the larger cases some of the rectangular structures are also performing better than the biggest square indicating that also for these dimensions there will be an optimum that will need to be mapped out for the specific devices.

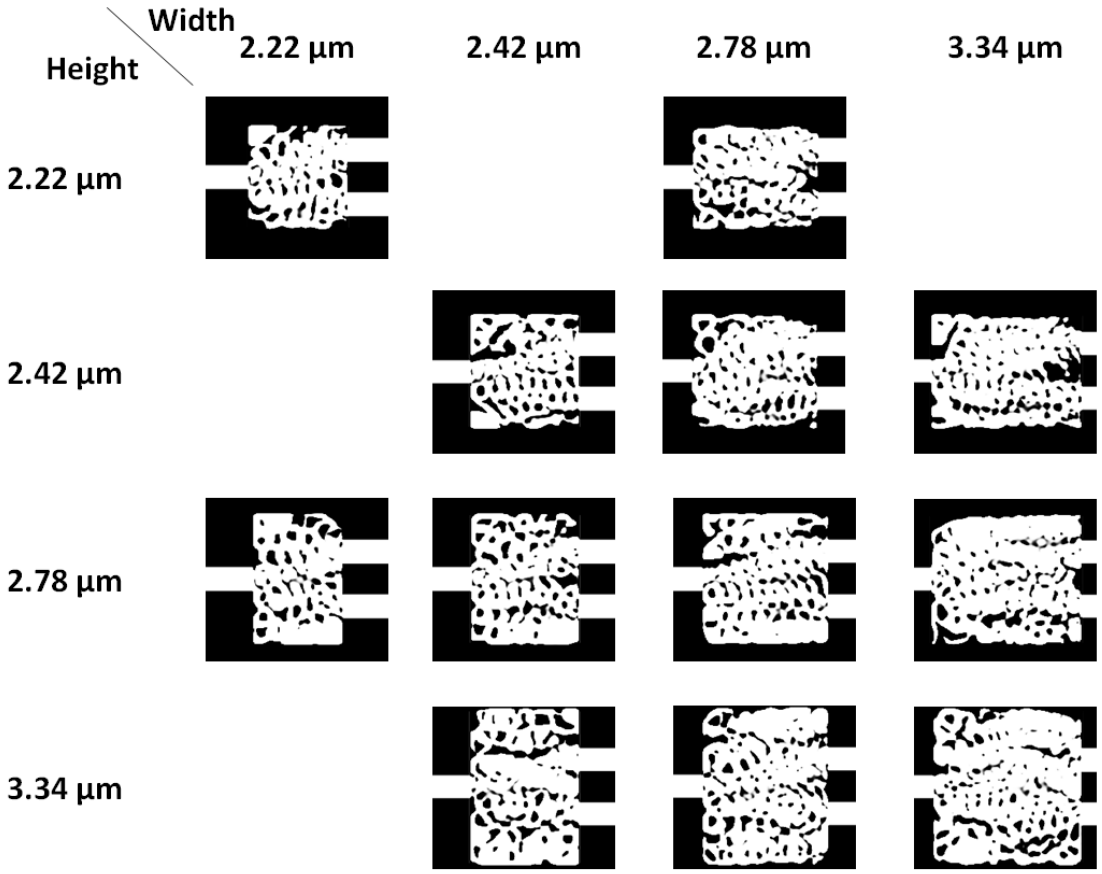


Figure 6.2.7: Topology optimized wavelength multiplexers obtained from rectangular design domains. The height and width of the sides have been indicated.

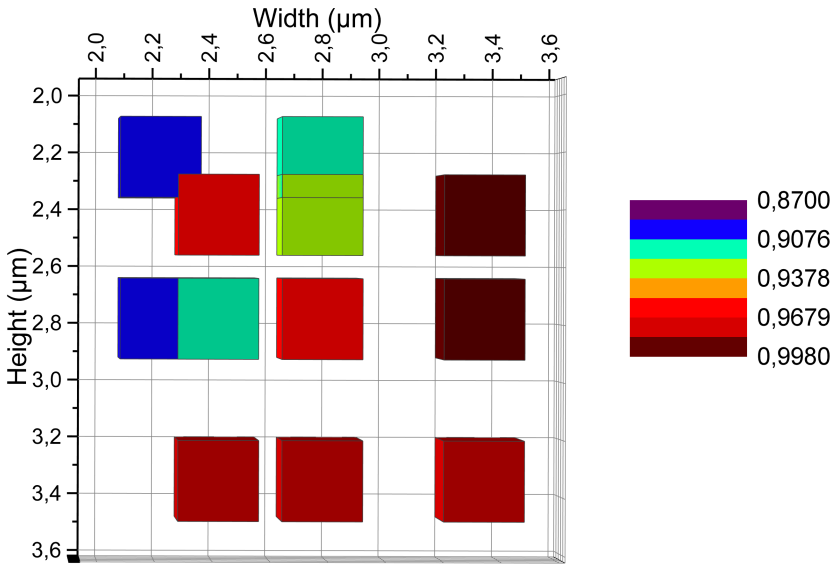


Figure 6.2.8: Objective functions presented in a colour scale plot as function of side lengths of the rectangular design domain corresponding to the designs presented in figure 6.2.7.

6.2.3 Asymmetric waveguides

For the well-functioning symmetric designs it seems that the objective in one output waveguide is more difficult to satisfy than the other. In the previous the summed objective function has been presented, however an objective function is obtained for each objective in the TO. As an example of a set of objective functions, including one for each of the two objectives in the demultiplexed waveguides and their sum, the results from the symmetric $2.78\,\mu\text{m} \times 2.78\,\mu\text{m}$ multiplexer is shown in figure 6.2.9(a). This example is representative for all of the functioning structures in that the 1550 nm objective function converges towards a substantially higher value than that of the 1300 nm objective. That means that this branch is the limiting factor and it is attempted to even them by making an asymmetric design where the input waveguide carrying the multiplexed signal is shifted towards the waveguide carrying the demultiplexed signal around 1550 nm. Based on the experience from the work presented in section 5.3 the waveguides are, however, not entirely aligned, the input carrying both wavelengths has imply been shifted downwards towards the height of the 1550 nm waveguide.

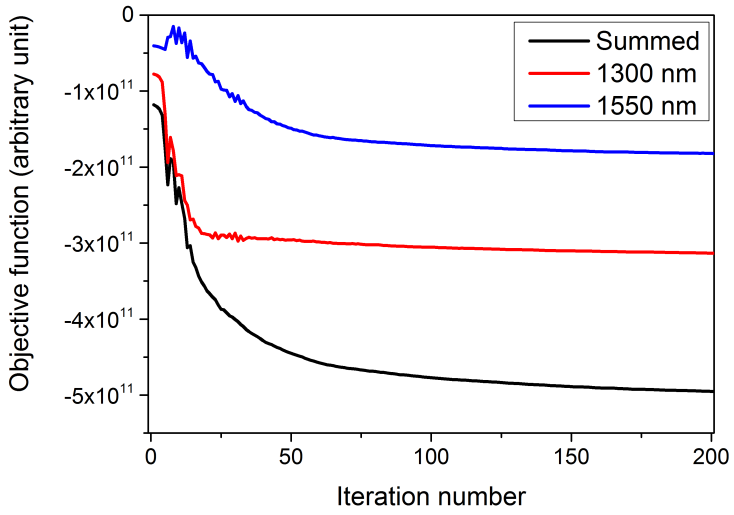


Figure 6.2.9: (a) Raw, unnormalized objective functions for the $2.78\,\mu\text{m} \times 2.78\,\mu\text{m}$ wavelength multiplexer shown in figure 6.2.4, including one for each of the two objectives as well as their sum.

Similar to the symmetric case, a number of different sizes of the square design domain is tested for this asymmetric configuration of the waveguides, a subset of the resulting structures are shown along with the normalized objective functions in figure 6.2.10; the results from the symmetric case are plotted along in grey for comparison.

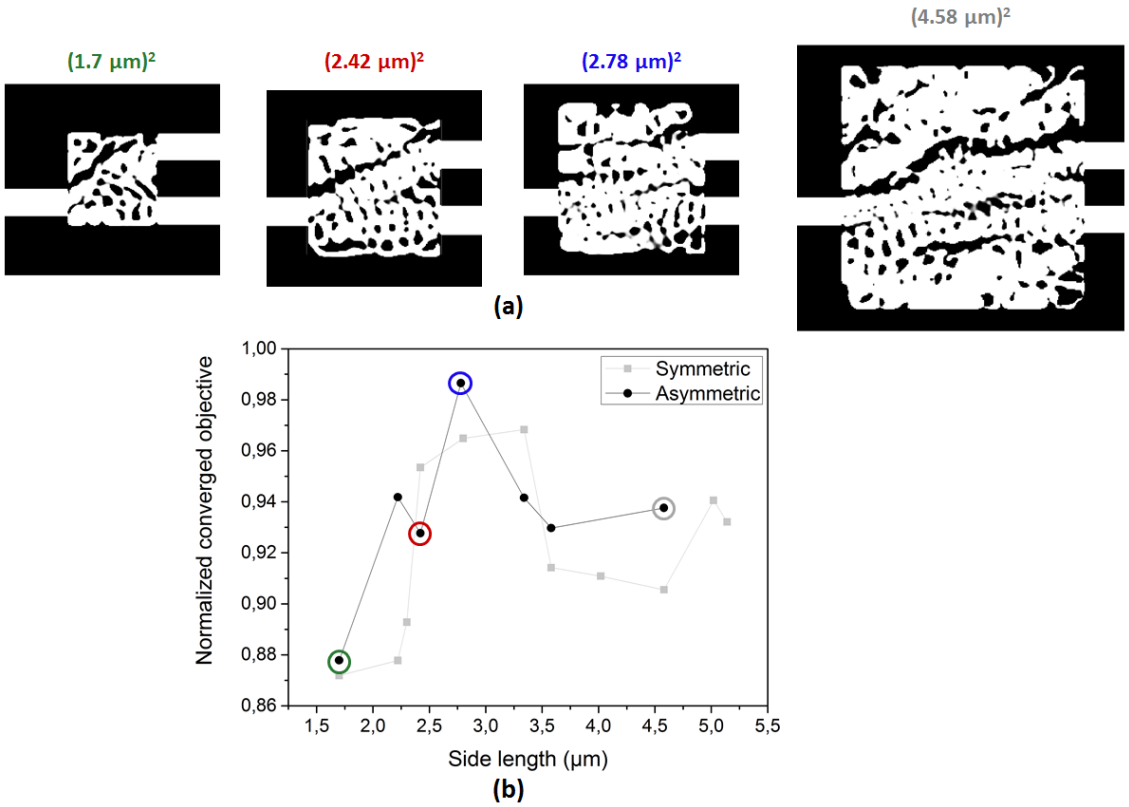


Figure 6.2.10: (a) Topology optimized designs for square wavelength multiplexers with an asymmetric waveguide configuration for various design domain footprints. (b) The normalized objective functions as a function of the side length. The connecting line is simply a guide for the eye as are the coloured rings marking the designs that have been illustrated. The results from the symmetric case is plotted in grey for comparison.

The final designs are quite different from the devices obtained from the square starting point structures as is to be expected. Similarly the correlation between side length and the convergence of the objective function is not the same, emphasizing that it is an analysis that must be performed for the specific device under consideration. In spite of these difference it seems that the overall level of convergence is higher for the asymmetric case including the best performing devices. This underlines that even though the TO method is strong, having an appropriate starting point can make a big difference as choosing too simple a structure for the optimization may limit the end results.

6.2.4 Waveguide spacing

Throughout all other work with the wavelength multiplexer the spacing of the demultiplexed waveguides has been 700 nm in accordance with the coupling analysis of figure 6.2.1. Meanwhile it is still of interest to see if a larger or smaller gap would lead to such a big difference in the device performance that subsequent movement of the output waveguides would be worthwhile. Simulations were made for square devices with a footprint of $(2.78\,\mu\text{m})^2$ and 5 different gap sizes of the vertically symmetric design, the results are presented in figure 6.2.11. The objective functions exhibit a clear trend of the smallest gap giving the best results and this trend is backed up by the spectra. However, there is relatively little difference for the smaller distances, and thus there is found no reason to lessen the distance below the 700 nm and needing to move them further from each other to avoid coupling after the demultiplexing.

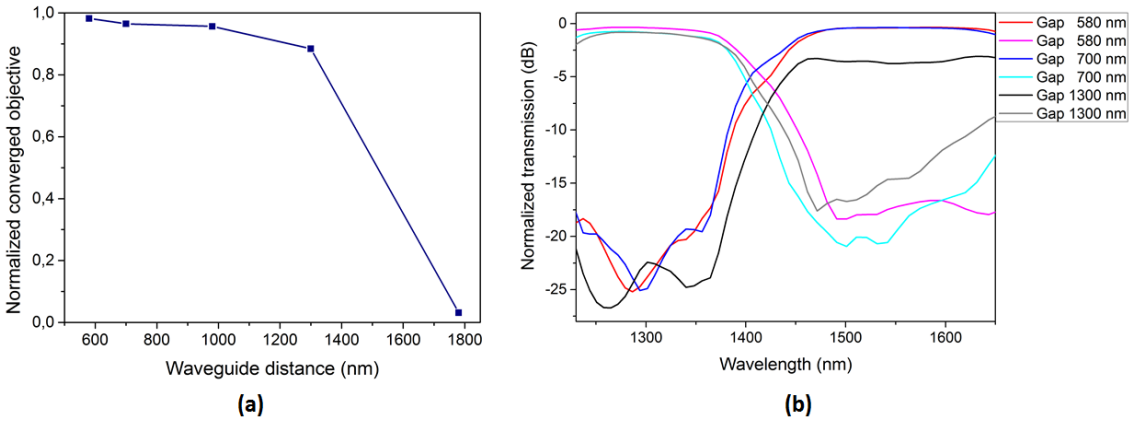


Figure 6.2.11: (a) Normalized objective function values at iteration 198 for the five different waveguide edge-to-edge spacings. (b) Comparison of the spectra for three of the waveguide spacings.

6.3 Feature size-performance-correlation

After the optimum design footprint has been determined for a specific device, there are other factors which need to be considered. Even for a relatively big structure there may be challenges in regards to fabrication, as the features within the structure will still me small. In this section the role of the feature size, as determined by the employed filtering, is investigated for the same type of square wavelength multiplexer devices as described in the previous section. Starting point structures are as those illustrated in figure 6.2.2.

6.3.1 Filter sizes

For the study of feature sizes a number of different DD dimensions were used. The resolution is 20 nm/pixel and the filters applied are 40 nm (same as presented in previous section), 80 nm, and 160 nm respectively. The resulting structures and the colour-mapped objective-based performance is illustrated for the symmetric structures in figure 6.3.1.

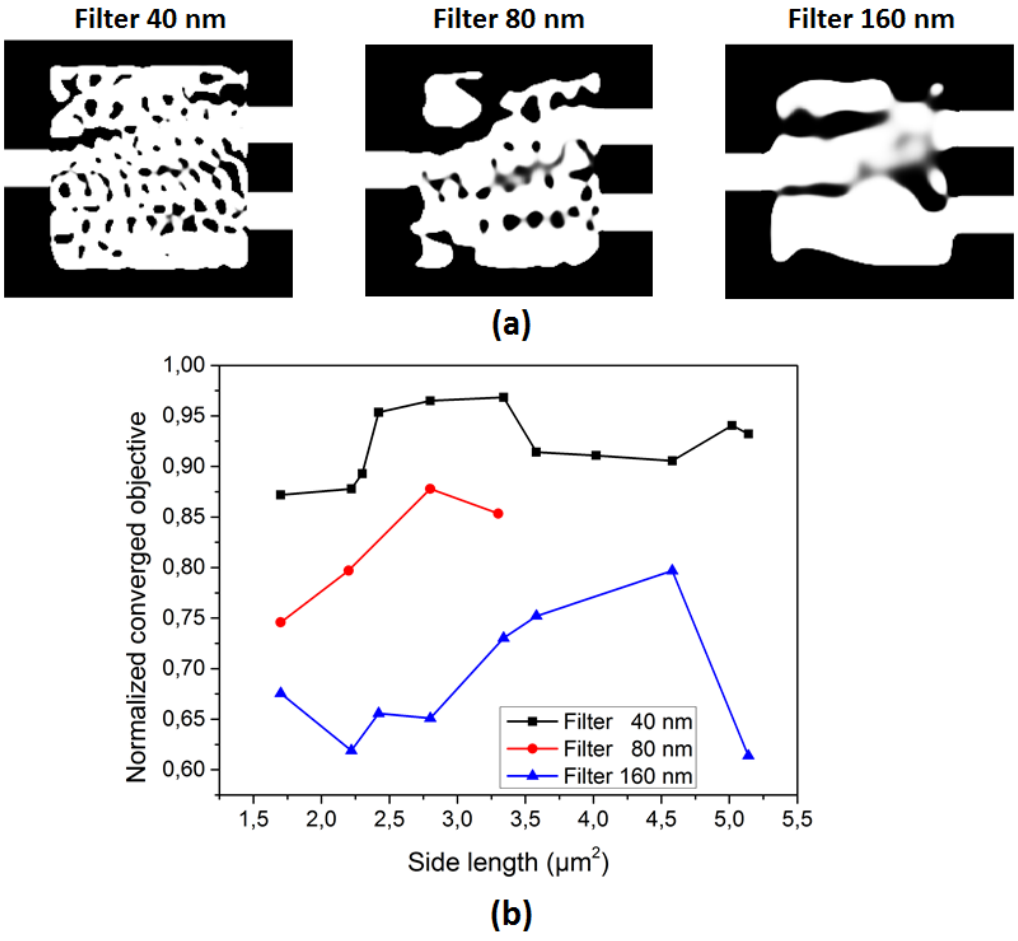


Figure 6.3.1: (a) Topology optimized wavelength multiplexers with symmetric waveguide positions and a footprint of 7.73 μm^2 optimized for three different filter sizes of 40 nm, 80 nm, and 160 nm respectively. (b) Normalized converged objective function value as function of design domain footprint shown for the three different filter radii.

Unfortunately, but not surprisingly, the smallest filter exhibits the best performance. The small features allow for more scattering and interference processes to occur within the small space and the lower filter gives the program more freedom. The optimum footprint size seems to shift upwards as the filter size increases, but

even the best performing device with the 160 nm has a worse performance than the lowest of the 40 nm designs. In figure 6.3.2 choice spectra are presented to illustrate the difference in performance of the various filters. Where the 80 nm simply has a lower extinction ratio and slightly lowered insertion loss it is apparent that too big a filter, as in the case of the 160 nm radii the functionality itself is lost.

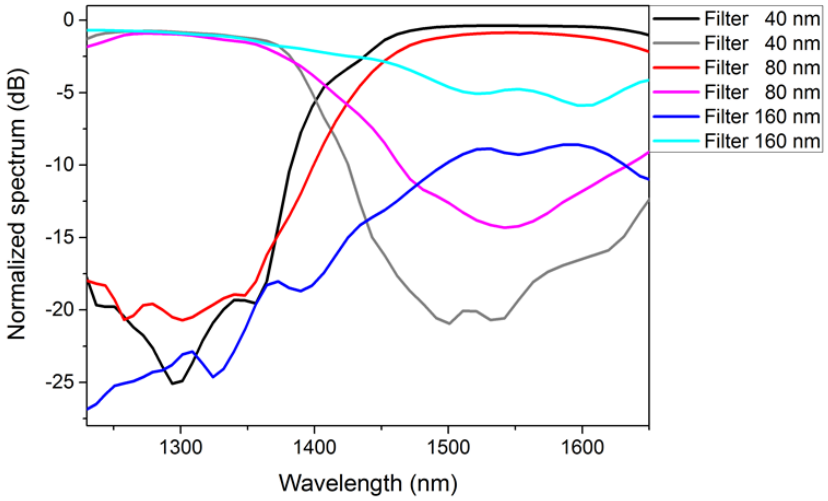


Figure 6.3.2: Spectra of the 11.16 μm^2 *footprint devices shown in figure 6.3.1 for the three different filter radii.*

This analysis shows that it is the smallest feature sizes and most complex structures that leads to the best design performances; this is also in correspondence with the experience from other designs and the investigation of the resolution which can be viewed as a coarse filtering. This is unfortunate as the smaller features are more difficulty in regards to fabrication, it is however a matter of trade-offs, one may choose to add a filter to make fabrication easier, however it will come at a cost of performance.

6.3.2 Reapplying filters

Although the larger filter sizes did lead to worse performance, they might still be helpful tools to reach more easily fabricated structures. In this section the results of reoptimizing structures, originally obtained with a large filter but applying a smaller filter during the reoptimization, are shown. For example the structure obtained with an 160 nm filter was used as a starting point for an optimization with the 40 nm filter. This is done with the settings of the optimization altered so that it allows for fewer grey areas in the beginning of the TO than usual. This is done to make sure that the new structure bears semblance to the starting point. In this manner the performance may be improved while not introducing as many very small structures. Furthermore the new starting point may lead to a different local minimum with better final performance.

In figure 6.3.3 the effects of reoptimizing the structures from the 160 nm with a 40 nm filter is shown and compared to the original devices and their functionality. It is apparent that reoptimizing using the smaller filter leads to increasing the functionality once more; in some cases even surpassing what was obtained from applying the 40 nm filter from the beginning. As intended the overall features of the high-filter structure are maintained and fewer small features appear in the reoptimized design. However, as is clear from the example in figure 6.3.3(a) that small features

will still be present although to a lesser extent. Similar trends are found if reoptimizing the 80 nm filter. However it does in general seem that for structures where the original optimization has been pressing the TO tool to its limits, so that the original performance was quite poor, reoptimizing will lead to a very different final design.

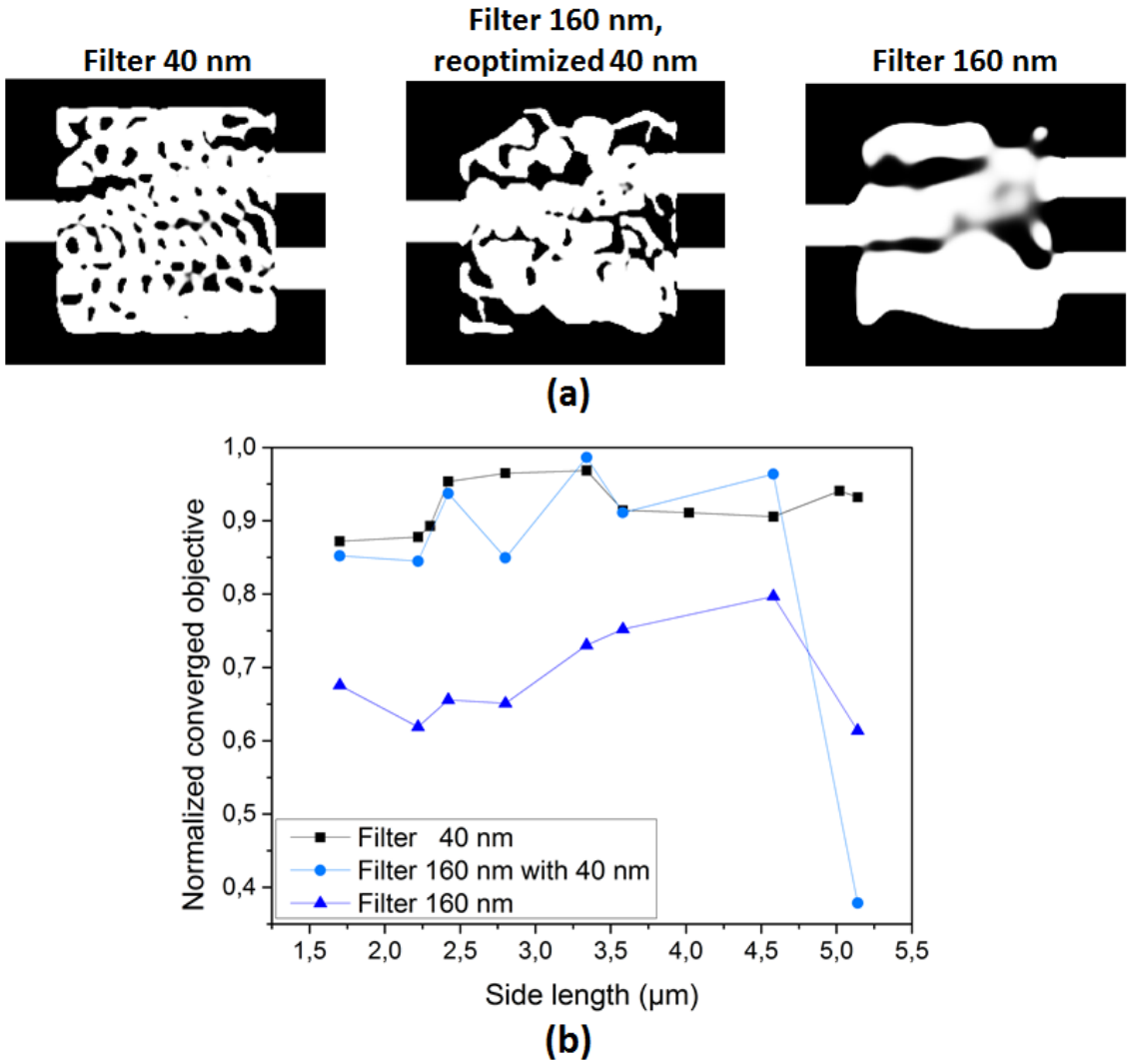


Figure 6.3.3: (a) Topology optimized wavelength multiplexers with symmetric waveguide positions optimized for different design domain footprint. Shown for an optimization with a 40 nm filter directly applied, with a 160 nm directly applied, and reoptimized with a 40 nm filter applied to the structures obtained from the 160 nm filter. All for a 11.16 μm^2 footprint device. (b) Normalized converged objective function value as function of design domain footprint shown for different filter configurations. "Filter 160 nm with 40 nm" refers to the optimization done first with a 160 nm filter and subsequently with the 40 nm filter, using the former as a starting point structure for the latter.

Similar analysis has been made to a smaller subset of the footprint sizes for the asymmetric starting point structure with the shifted waveguide for the multiplexed signal. The results are summed up and illustrated in figure 6.3.4; in general the conclusions are the same as for the case of the symmetric structure. Structures with smaller filter radii, and thus many small features, perform best. Structures that may be easier to fabricate due to less complicated features yet exhibiting high functionality may be achieved by using structures where a high filter size was applied as starting point structures for reoptimization with a smaller filter size.

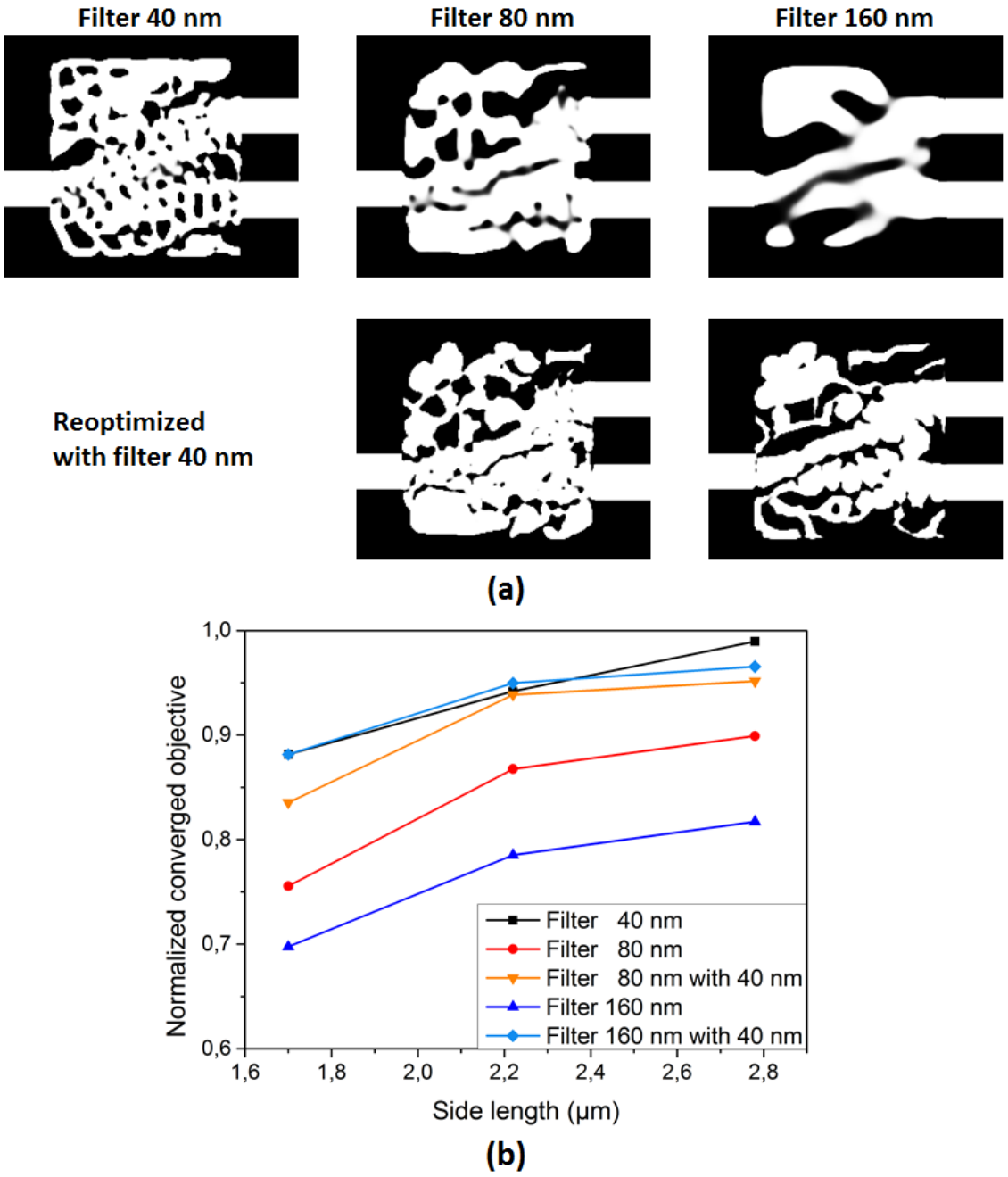


Figure 6.3.4: (a) Square design domain topology optimized wavelength multiplexers with asymmetric waveguides and a side length of 2.78 µm for various filtering configurations. Top row has three different filter radii while the bottom row shows reoptimization using the top row structures as a starting point and applying a 40 nm filter in the end. (b) Comparison of the performance quantified by the normalized converged objective value for the asymmetric multiplexer with various filter configurations.

6.4 Comparison to other studies

Although what is presented here will not be a direct comparison it would not be prudent to disregard that Pigott et al. have recently published similar work. To put the size and feature analysis in perspective it is attempted to recreate the results published in [22] and compare them to the structures designed using TO.

Firstly it should be mentioned that Piggott et al. use a different set of material parameters than has otherwise been employed in this project[4]. In this section however the following is employed:
$\epsilon_{Si} = 12.180$
$\epsilon_{SiO_2} = 2.1025$.

The design domain is square with side lengths of $2.78\,\mu$m, however the design differs from what is earlier presented as the distance between the waveguides carrying the demultiplexed signal is 585 nm. As per section 6.2.4 the smaller spacing is expected to give better results but it will lead to more coupling between the waveguide which means that the spacing must be increased after the de-multiplexing has taken place.

For this analysis the design presented in [22] has been imported as a cut-out from the pdf of the published paper, the resulting structure is shown in figure 6.4.1 (top left). This procedure of extracting the structure from the paper is one reason that this cannot be considered a fair comparison as it will introduce changes in the structure. However the design is evidently still present and as can be seen from the spectrum in figure 6.4.2 the functionality is maintained. The insertion losses for our FDTD calculations are also very similar to the $\approx -2\,$dB reported in [22].

Figure 6.4.1 shows the imported objective first structure when various filter sizes have been applied and corresponding topology optimized designs using the former as input structures. Applying the larger filters muddles the structure which is imported, but in all cases the topology optimized structures bear semblance to the original structure. However the structures have clearly been modified. In the small filter cases it is especially clear that more small structures have been added. This is likely to enhance the performance as is also apparent from the corresponding spectra in figure 6.4.2. This is however another reason why this cannot be considered a direct comparison of the two methods: The paper states nothing on the use of filters, but evidently some have been applied; however it is not possible to recreate properly the existing conditions. Applying a 60 nm filter for the topology optimization reveals the structure with features most like the original objective first design yet it has clearly enhances the performance as seen from the insertion losses in figure 6.4.2(a) (green). The differences between the various filters applied to the TO has only a marginal difference on the insertion loss and also relatively low on the extinction ratio.

The bottom row of figure 6.4.1 shows designs obtained from TO applied to a simple box structure but rather than the objective first inputs, but with all other parameters the same. It is evident that the TO tends towards a different local optimum when not given an already functional starting point, comparing the spectra of figure 6.4.2(b) this emphasises that there are many different local optima and that their performances are very similar. Comparing this structure to the $(2.78\,\mu\text{m})^2$ device in figure 6.2.4 even though the specific features are same the overall structure is definitely recognizable in spite of the changed parameters indicating a tendency towards this type of optimum for the TO program.

[4]Noted in chapter 2 and reiterated to be $\epsilon_{Si} = 11.68$ and $\epsilon_{SiO_2} = 2.082$ for easy comparison.

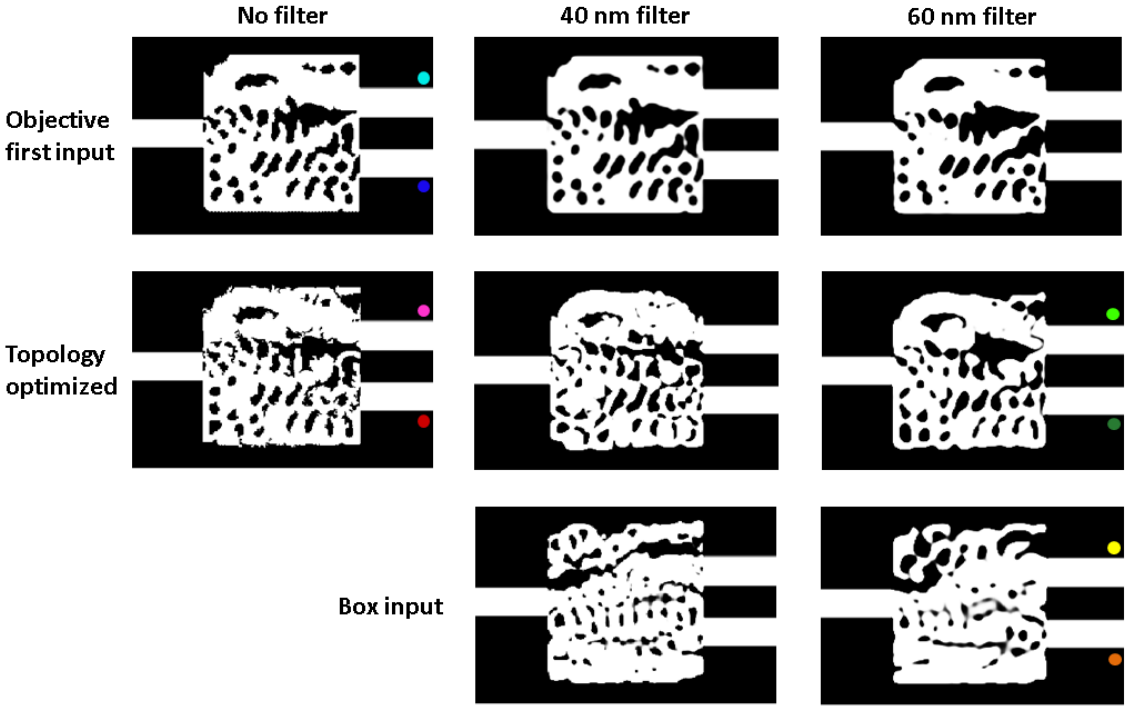


Figure 6.4.1: Optimized structure for the wavelength multiplexer. Top row shows the devices obtained from the objective first method [22] after it has been imported into Phazor with different filter sizes. Mid row are the structures gained from topology optimization using the above structure as a starting point. Finally the bottom row designs was obtained from applying topology optimization to a simple box structure all other parameters being the same as used by Piggott et al. The coloured spots added in the corner are for easier reference with the spectra presented in figure 6.4.2.

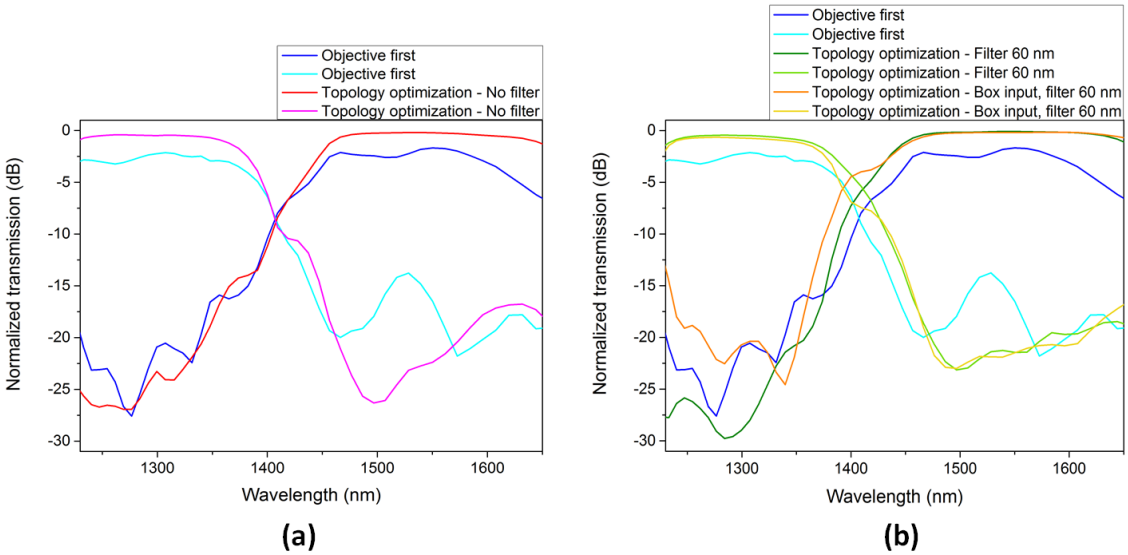


Figure 6.4.2: Simulated spectra of the wavelength multiplexers based on and in comparison with the work from [22]. The simulations through the imported design from Piggott et al. has been included in both graphs for easy comparison. (a) Spectra of the directly imported structure and the spectrum obtained from applying TO to that same device. Neither has included a filter. (b) Comparing the performance of the devices obtained from a square starting point structure and the inverse design structure as a starting point, respectively.

As has been emphasised the representation of the objective first structure as simulated through Phazor will be degraded due to the handling of the image and the potentially larger filter of the optimization may also be a limiting factor of the performance of the objective first value. Thus it cannot be concluded if either method is better in general terms, however it is safe to say that they are both useful tools for creating compact nanophotonic devices and that it is likely that both will benefit from intelligently chosen starting points as per the analysis of this chapter.

6.5 Summary and discussion

In this chapter, the effects of footprint and features sizes have been under systematic investigation for the specific case of a wavelength multiplexer. It was found that once a certain size has been reach it is unlikely that increasing it further will improve the performance significantly. This range is however individual for the specific device and is sensitive to the fluctuations in convergence to local minima. Thus this type of analysis must be done for the specific device under investigation to obtain the most compact high functionality design. Meanwhile a general, and interesting, conclusion from this is that increasing the design domain will not always lead to a large performance benefit.

The analysis of the feature size led to the more general conclusion that smaller features, as obtained by using smaller filter radii, will lead to a better functionality including lower insertion losses and cross-talk. This is unfortunate as smaller feature sizes are more cumbersome during fabrication. Using a large filter structure as a starting point for reoptimization with a smaller filter, and restricting the topology optimization to less change, may however lead to new designs with fewer small features as a compromise solution. This is a general lesson: that a more refined starting point structure accommodating of the functionality which is desired can lead to better devices. These results are exemplified in figure 6.5.1, which shows the spectra of two optimized devices with similar functionality. One has used a simple square, symmetric starting point similar to the one used in [22], the other has been given a more refined asymmetric starting point and has been optimized twice with two different filter sizes. By this additional treatment of the process for the latter structure, it has been possible to reduce the footprint by $2.8\,\mu\text{m}^2$. The trends appearing from this analysis may of course be skewed due to structures converging to different local optima, however as there currently exist no inverse design method ensuring that the global optimum is found, it is still highly relevant. Taking the extra steps during optimization is not necessary to show that a design type is functional or can be achieved by inverse design methods, however they should be employed when aiming for optimum devices to be integrated in large scale systems.

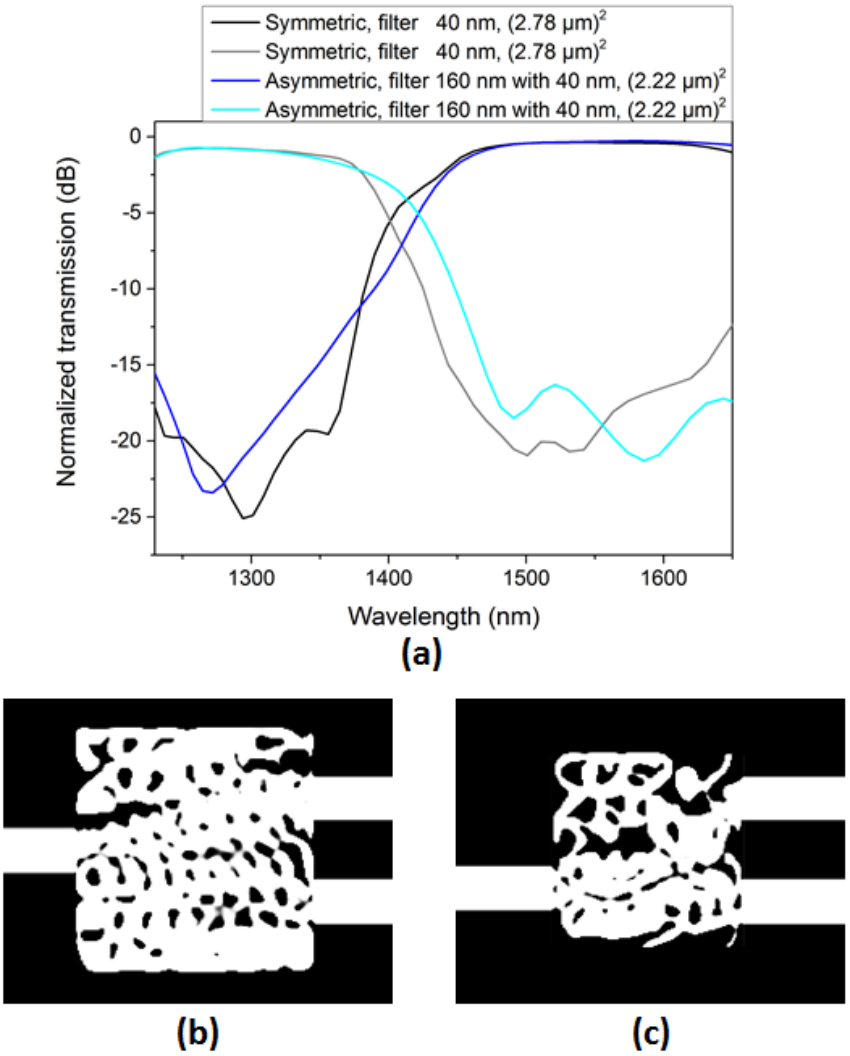


Figure 6.5.1: Comparison of the structures obtained for a simple and a refined starting point. (a) Spectra of the two structures revealing approximately the same functionality. (b) Device optimized from a vertically symmetric starting point structure with side lengths of 2.78 μm. (c) Device obtained with an asymmetric starting point structure where one waveguide is shifted. Side wall lengths of the square design domain are 2.22 μm and an optimization with a 40 nm filter has been applied subsequent to one with a 160 nm filter.

CHAPTER 7

COMPACT TAPER

As has already been highlighted in this thesis, the small footprint is an important feature of devices for PICs. Due to the strong confinement in the SOI platform, the silicon waveguides usual have a width of less than 0.5 μm. Nonetheless there are a number of functionalities that requires devices with a larger size. Various devices are likely to be designed with different cross-sections, for example grating couplers typically have dimensions around 12 μm. Because of such size differences it is important to be able to change the spot size on the chip while undergoing a minimum of conversion to higher higher-order or radiation modes. This is a functionality that can be achieved with a relatively simple design of a linear adiabatic taper; however this takes op a lot of space. Taking the example of converting from a single mode 0.5 μm photonic wire to a 12 μm waveguide, the conventional linear taper must be longer than 300 μm [61]. Numerous methods have been employed to achieve low-loss tapers with compact footprints, however there tends to be a trade-off of size and loss for these devices as well.

The methods that have been used for making compact tapers can be put into the following three categories which have been illustrated by concept sketches in figure 7.0.1.

Adiabatic tapers [61, 62] are build on the principle of gently and monotonically widening the taper. If the increase in waveguide width is slow enough it will not impact the propagating field. Adiabatic tapers includes the linear, parabolic, exponential, and Gaussian taper types. For the linear taper the limitation, that ensure that there is no mode conversion, is that the spreading of the waveguide sidewalls does not exceed the diffraction spreading of the lowest order mode. The performance achieved for the optimized adiabatic taper of [61] is a transmission of about 98 % for a taper length of 120 μm connecting input and output waveguides of 0.5 μm and 12 μm.

Non-adiabatic planer waveguide tapers [63, 64] use modulations of the sidewalls of the tapering region to influence the field whereby a more compact structure, than could be allowed by a linear taper, is made possible. Various methods exist where the widths are modulated either in square segments or by points along a curve forming the shape of the taper edge. The structures are always symmetric along the direction of propagation. The designs are formed through genetic optimization procedures and tend to converge close to the global optimum. The rectangular segmentation structures show better functionality, however they are difficult to realize due to rounding during fabrication and these designs are yet to be experimentally

verified. For a 1 μm taper connection a 0.50 μm photonic wire with a 2.0 μm waveguide theoretical transmission efficiencies of 85 % have been achieved using for the non-adiabatic tapers. They show promising results for compact structures, but the sizes seem to be limited by the long computational time.

Lens-assisted focusing tapers [65] have a double concave lens built into the wide end of the waveguide and the length of the taper is then determined by the focal length of the lens structure. In practice the lens is created by etching the desired geometry into a region of the silicon waveguide thus creating a thinner silicon slab and a lower effective refractive index. Connecting waveguides with width of 0.45 μm and 10 μm with tapers of length between 10 μm and 20 μm has been done and experimentally verified with transmission of around 80 % for TE-like polarization. It allows for a substantial decrease in length compared to the adiabatic case, however at the cost of a higher loss.

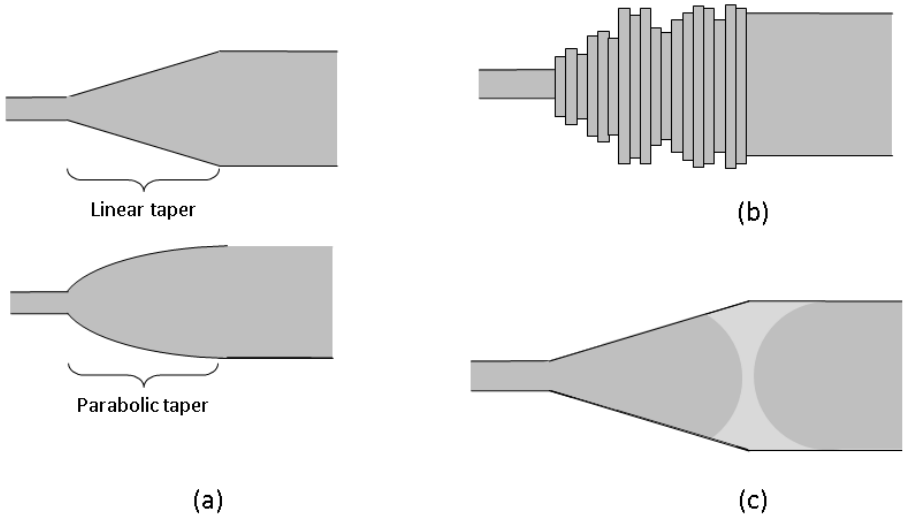


Figure 7.0.1: Sketches of various principles used for compact tapering. (a) Adiabatic tapers, simplest principle of the linear taper as well as the slightly refined parabolic taper. (b) Nonadiabatic taper where the field is modulated by section-wise variation of the taper width. (c) Lens-assisted focusing taper. A lens is created at the wide end of the taper as the effective refractive index is altered by different thicknesses of the silicon layers.

In this chapter attempts of creating a compact taper using TO is presented. A conceptual sketch of the starting point structure is shown in figure 7.0.2. In all cases the narrow waveguide had a width of approximately 0.440 μm (may be altered slightly when changing the resolution), however the width of the wide waveguide was altered throughout the different design iterations. The length of the tapered region was altered throughout in an attempt to find the smallest length revealing acceptably low losses.

In this chapter the issue of exciting the fundamental mode in the wide waveguide will be addressed first. Then various design configurations obtained with different input strategies and found in both 2D and 3D will be presented. Lastly the optimization of smaller structures of no more than a few μm will be introduced along with a different, step-wise, optimization scheme. Finally a summary along with a discussion of the approaches and future work on the design concludes the chapter.

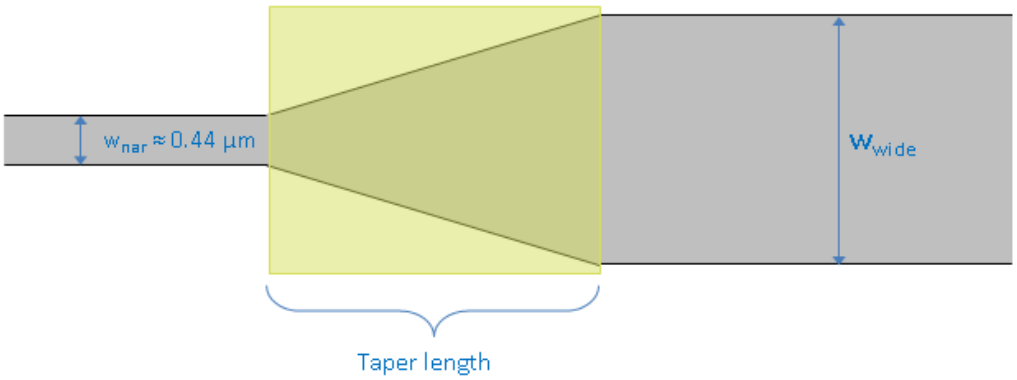


Figure 7.0.2: Sketch of the basis structure for the optimization of a compact taper. The narrow waveguide has a width of $0.440\,\mu m$ *while the width of the wide was varied to be around* $6\,\mu m$ *and* $12\,\mu m$ *respectively. In either case the design domain (yellow) covers the linearly tapered region the length of which was also changed throughout the various optimizations.*

7.1 The wide fundamental mode

The biggest challenge in regards to this device optimization is the great width of the waveguide carrying the fundamental mode. As the current implementation of Phazor does not have a mode exciter or objective, it becomes difficult to operate with the fundamental mode on large scales. Another issue making this even more challenging is that the calculation time naturally increases as the sizes increase, limiting the possibility of various workarounds. The calculation time for this device is already relatively high due to the wide waveguide.

For this optimization problem there are two possible configurations: Either the source is in the wide waveguide and the objective in the narrow or vice versa. Both have been investigated, but since it is easier to evaluate the functionality of the source through simulation alone, preference has initially been given to exciting the fundamental mode in the wide region. Here the various attempts will be presented and compared.

The initial work was done in 2D to get an overview while keeping the calculation times low. Although 2D is unlikely to lead to optimum structures for fabrication, it is evaluated that a starting configuration would not work in 3D if it does not in 2D; thus the 2D test is still worthwhile. In figure 7.1.1 the most straight forward source configurations are shown in a straight waveguide. This includes a single source with the same width as the waveguide, one which is shorter and one which extends beyond it. The full width shows the best performance, even so interference is clearly present and a TE_0 mode is not maintained. Curiously a similar effect is seen to take place for the case of the short source the light is relatively confined to that width of the exciter within the waveguide, yet similar interference patterns arise. Extending the source beyond the waveguide in the simulation has very little impact.

Similar to the method that was used for the TE_0 mode when creating the 3-mode multiplexer in section 5.4, it was attempted to use a tapered input as shown in figure 7.1.2. Although this configuration is a potential solution, it has some severe practical limitations. The main issue of this method is the increase in the image size due to the leading taper causing a great increasing the optimization time. This poses a limit on the length of the input taper and it will become even smaller in the

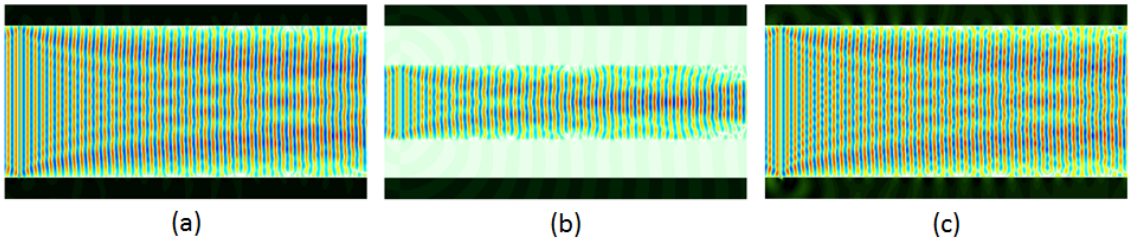


Figure 7.1.1: Attempts at exciting the TE_0 mode through placement of a single source with a uniform intensity distribution in the wide leading region of a tapering structure. (a) A single source extending through the entire width of the waveguide. (b) A single source with a width half that of the waveguide. (c) A single source which extends beyond the edges of the waveguide.

case of 3D optimization. It is thus not possible to use an adiabatic taper exceeding a length of 100 μm as is necessary for low loss fundamental mode operation. A higher loss configuration would however not harm the optimization procedure as such as it simply maximizes the output. In addition to the problem of size figure 7.1.2(c) reveals the issue of the reflectance occurring when the broad TE_0 mode is incident on the short tapered region which is to be optimized. This leads to strong interference and ruins the TE_0 mode. This effect is an issue for all instances where the objective is placed in the narrow waveguide; but strongest in the case of the long input taper where it is not possible to include a long straight region as well and where further resonance can occur due to the angled sidewalls of the leading taper.

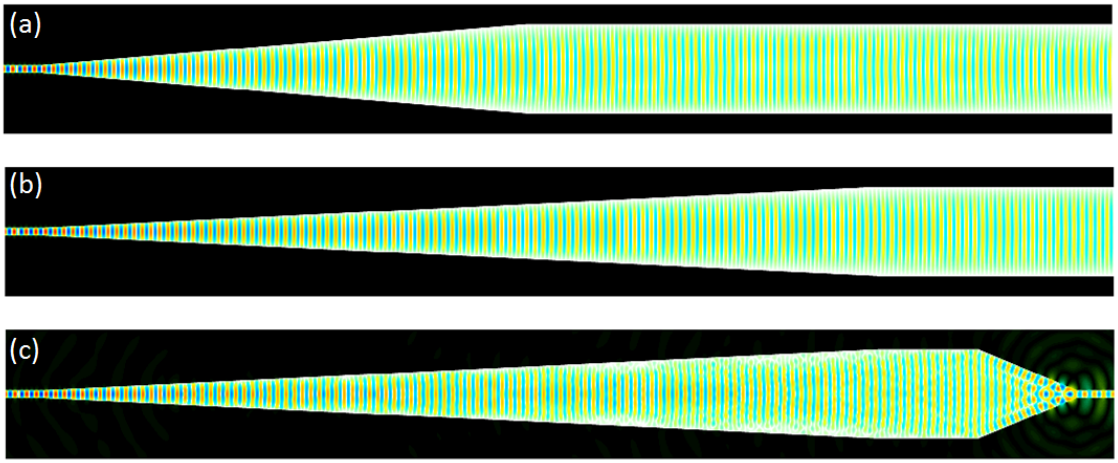


Figure 7.1.2: Illustration of the use a long tapered region as input for the wide waveguide to obtain a wide fundamental mode as an input for the optimization. Structures overlaid with the 2D FDTD calculated H_z field. (a) 30 μm tapered region, (b) 50 μm, and (c) a 50 μm tapered region for the input followed by a 6 μm taper which is to be optimized.

The last method employed to excite the TE_0 mode in the wide waveguide is to manually space the source points in the pixels running across the waveguide to emulate the mode shape. This was attempted both in a heuristic manner including sources with a width of several pixels and with spacing regions including no exciters in between allowing for their interference to take part in shaping the pulse. Although some of these configurations come close to a fundamental mode shape, all attempted configurations exhibited 'shoulders' indicating coupling to other modes as well as unwanted interference. The potential benefit of this method is that it can be adjusted to account for the reflections of the tapered region which is to be optimized. Unfortunately the reflections are changed during the optimization pro-

cedure and this leads to different resonances, rendering it impossible to completely cancel out this effect.

In a more stringent approach the field distribution of the TE_0 mode was first simulated using Lumerical's Mode Solutions. From this an intensity profile can be extracted and an exciter can be placed in every single pixel point across the waveguide with their amplitude adjusted in accordance withthe intensity of the mode profile; figure 7.1.3 shows the results of exciting the source in this fashion. Several iterations were made, firstly the mode profile was interpolated to fit the right pixel number, subsequently manual adjustments were made to ensure monotony and symmetry of the profile. Albeit not perfect this is the solution that has led to the best mode profile in the wide waveguide.

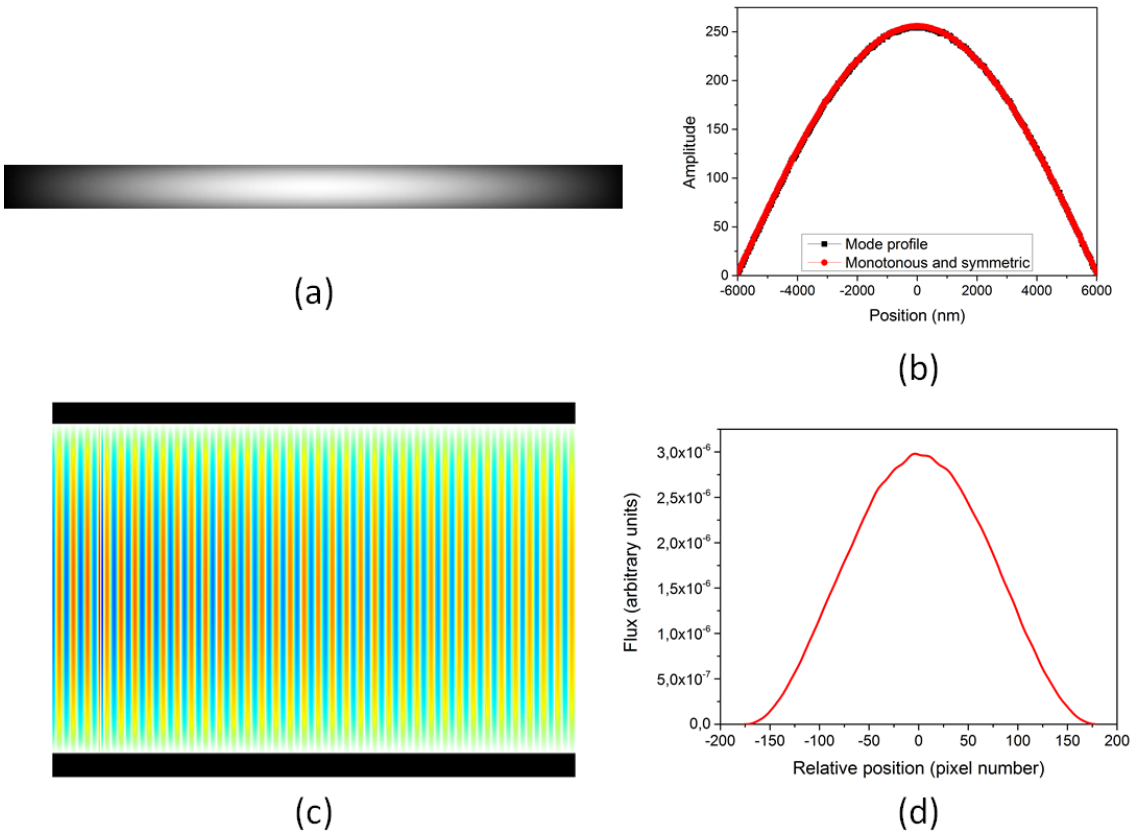


Figure 7.1.3: (a) Cross-section of the TE_0 mode profile in a 12 μm wide waveguide simulated using Lumerical Mode Solutions. (b) The extracted intensity distribution along the center line of the mode profile in (a) along with a profile which has manually been altered to be monotonous and symmetric. (c) Field simulation through a straight wide waveguide, obtained from a exciters placed in the 301 pixels along the width of the waveguide and with amplitudes adjusted in accordance with (b). (d) The flux profile of the simulated field through the straight waveguide; although some noise is present it does bear resemblance to the original profile in (b).

7.2 Optimized designs

Numerous attempts were made at optimizing compact tapers using the various types of sources described in the previous section as well as exciting the light in the narrow waveguide and optimizing for the wide output. In this section the resulting structures will be presented and their performances compared.

7.2.1 Tapered input

A number of optimization attempts were made using the long tapered input to gain the wide fundamental mode. In this case the starting point structure had a 30 μm leading taper between a narrow 0.480 μm taper and a broad 5.536 μm taper. The TO was performed on a 4.0 μm taper between the same waveguide widths. The structure was promising as shown in figure 7.2.1(a) and it was fabricated following the procedure described in chapter 3. Unfortunately the measurement revealed that the device did not in fact produce a TE_0 mode experiencing mild resonance but rather a noisy higher order mode. The approach of the long tapered input preceding the optimized region was abandoned.

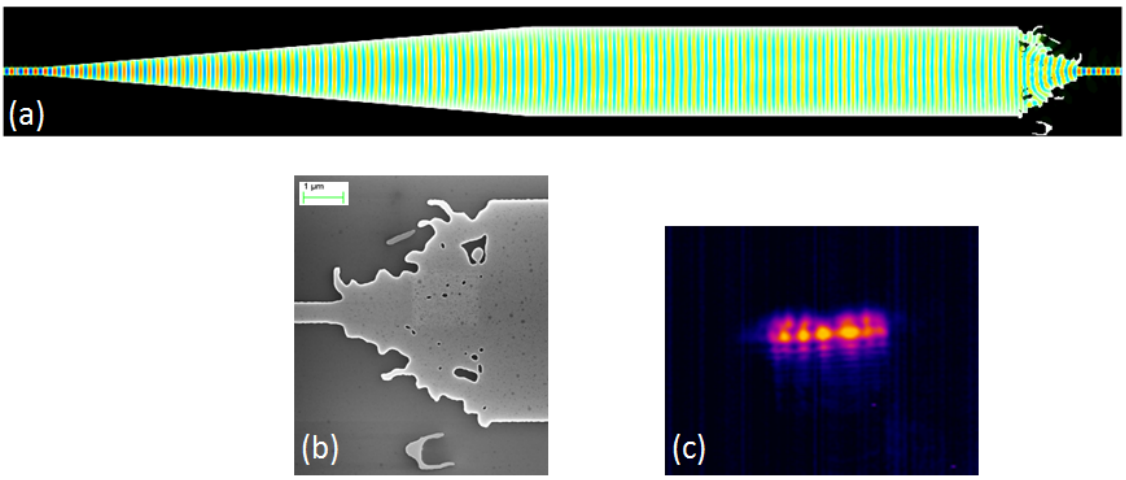


Figure 7.2.1: (a) Compact taper obtained from topology optimization on a 4.0 μm *region. A long leading taper has been introduced to obtain the* TE_0 *mode for the optimization. (b) Close up SEM image of the optimized region. (c) Mode profile of the light input in the narrow photonic wire and having traversed the optimized taper recorded at* 1570 *nm.*

7.2.2 2D investigation of taper length

Previous work has revealed that the full 3D optimization is important for the final performance. Yet there is a chance that the broadband functionality will lead to some correspondence between the 2D and 3D designs. Furthermore the 2D analysis is very relevant in this case where the input image is very large leading to long calculation times for optimization. Presently the necessary length of the optimized tapered region is investigated in 2D. This is worthwhile under the hypothesis that it will not be possible to reach a functioning design in 3D for dimensions smaller than those achievable in 2D. All devices were optimized for an objective placed in the narrow waveguide and the monotonous, symmetric exciter from figure 7.1.3 placed at the same position in the wide waveguide. The width of a rectangular DD was fixed at 12.6 μm and length varied to match the length of the taper.

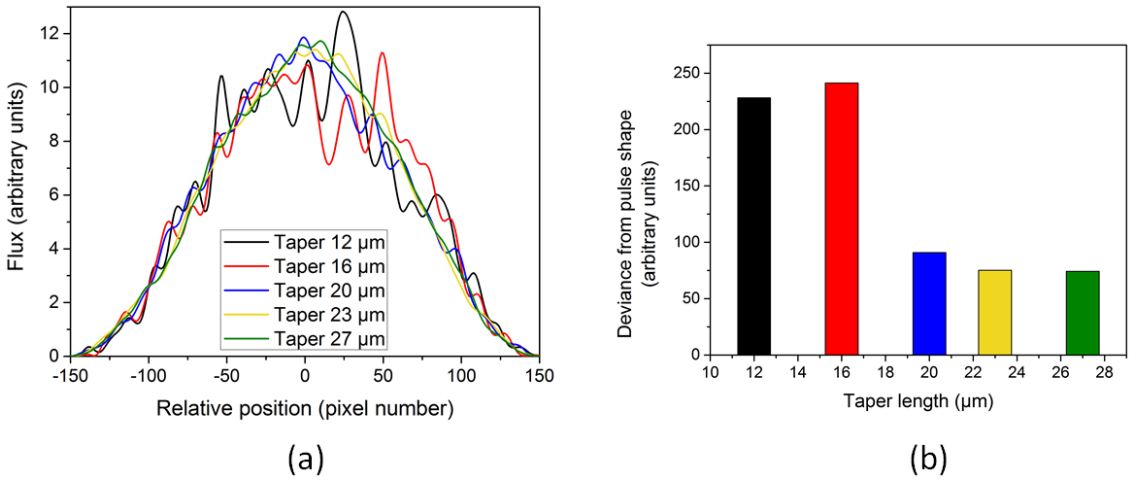


Figure 7.2.2: (a) Comparison of the simulated flux after optimization presented for numerous different tapering lengths. The design domain has the same width of $12.6\,\mu m$ *in all cases and the same length as the taper. During optimization the source was excited in the wide waveguide using the monotonous and symmetric profile exciter of figure 7.1.3. (b) A visualization of the deviation of the flux profiles of the optimized structures from that of the excited pulse. The plotted values are the integrals of the absolute values of the difference between the optimized and the intended pulse shape.*

It is apparent from figure 7.2.2 that too short a taper will result in a very noisy signal. For the tapers that are $\geq 20\,\mu m$ the profile shape is still noisy but it bears a clear resemblance to the sought TE_0 mode. In figure 7.2.3 the fields and flux of the shortest taper that exhibited a TE_0-like profile is shown. The field simulations were conducted with both 2D and 3D FDTD for comparison. It is apparent, especially from the flux profile, that the performance deteriorates noticeably when changing to 3D simulations of the structure optimized in 2D. This is not unexpected given the previous work but is still tested to confirm that it is the case for the big yet relatively simple structure of the taper as well. Furthermore this 2D analysis has led to the knowledge that the taper should be $\geq 20\,\mu m$ to function; this might be different in 3D but it is unlikely to be possible to make a taper much shorter, so this gives a starting point to work from. Additionally some semblance of the desired mode shape is seen even for the 3D FDTD and thus the 2D optimized structure can be used as a starting point for the further optimization to guide towards a specific local optimum. Unfortunately the 3D optimization based on the 2D knowledge did not in fact lead to a fully functional device. The failure of this method of exciting the TE_0 mode using sources placed in each pixel across the waveguide when optimizing in 3D may be due to the mode not spreading properly in the z-direction. Currently only a single row of sources is placed and further work could be done creating a plane of sources with different distributions in the various rows across the height of the waveguide.

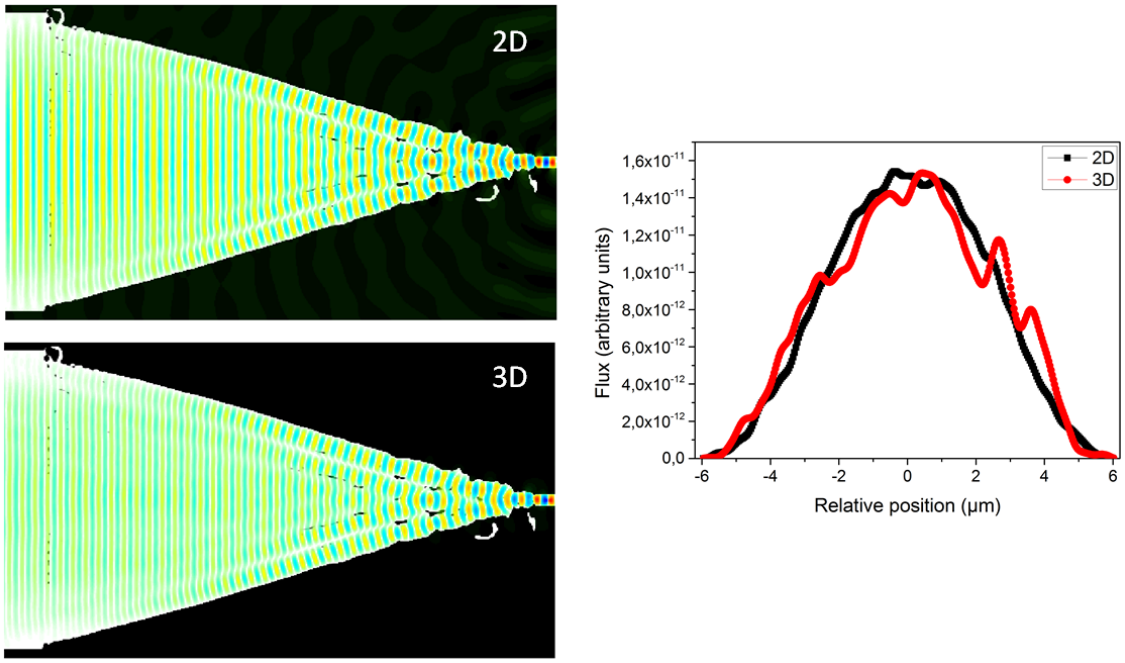


Figure 7.2.3: 2D optimized design for a 20 µm long taper with the field simulated in 2D and 3D respectively. The simulated flux after conversion is shown for a direct comparison of the devices.

7.2.3 3D optimized tapers with swapped source and objective

As none of the above techniques produced a desired outcome the reverse approach was attempted. In this case the source is placed in the narrow waveguide and the optimization objectives in the wide region. Depending on the objective either the weights will be adjusted to fit the mode profile (for the point energy objective) or the amplitudes of the single points are altered while the weights all have the same value (the pulse shape objective). In both cases the set-up is made so that there is one objective in each pixel perpendicular to the propagation direction of the waveguide.

Unfortunately neither method led to usable devices. Figure 7.2.4 shows the structures and field simulations for three devices with a 27 µm long tapered region optimized using the PSO. The only difference between the designs is the arrival time of the pulse, in figure 7.2.4(b) the pulse is to arrive 89 fs later than in (a). Although neither has the desired function, it is apparent that simply changing the time has a significant impact. Firstly for the longer traversing time, more scattering occurs. Secondly both cases show a clear tendency towards 'bending' of the field in the wide region. This is presumed to occur as the distance that the light must travel to reach the edges of the wide waveguide is greater than needed to reach the center. Even though the wave front is meant to arrive simultaneously at any point in the cross-section of the wide waveguide, it might be that the optimization can be guided further by introducing a time delay at the center. The structure resulting from such a tuning is shown in figure 7.2.4(c). Introducing this artificial time delay does seem to spread the light more into the wide waveguide than when using the central arrival time only. Unfortunately properly adjusting the arrival time of every single pixel is not a feasible solution and the approach of optimizing for the wide waveguide region was abandoned.

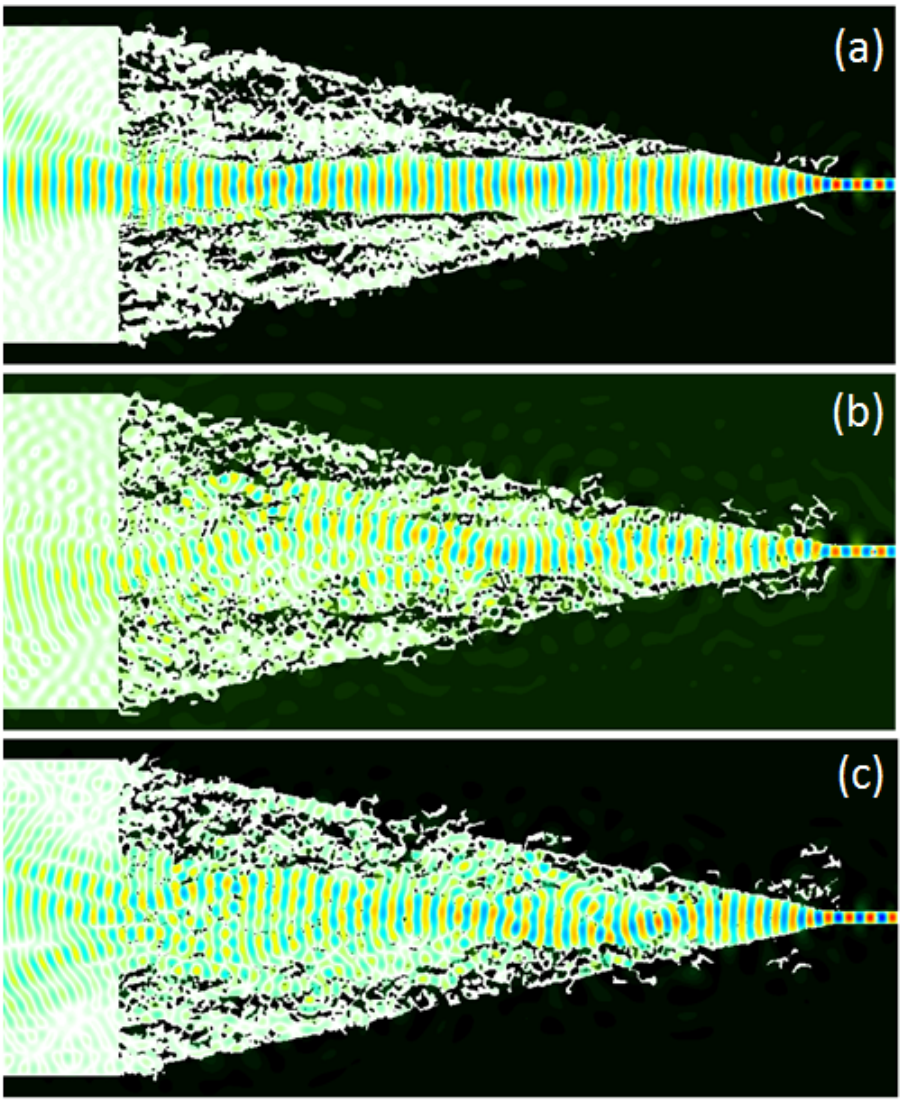


Figure 7.2.4: 3D optimized designs of a $27\,\mu m$ *long taper along with* H_z *field simulations. The objective is placed int he wide waveguide and are of the pulse shape form. (a) Optimized design obtained from a propagation time from source to objective of* $411\,fs$*, corresponding to the time from center to center in the unoptimized structures. (b) Increased delay time from source to objective,* $500\,fs$*. (c) Design gained by having a non-constant time delay across the wide waveguide, introducing a delay at the center.*

7.2.4 Alternative optimization attempts

A few alternative DDs and structures were attempted due to the difficulties with the excitation of the wide TE_0 mode, these will briefly be mentioned here for completeness.

Splitting the DD so that an unaffected region with the width of the narrow waveguide persisted was attempted in combination with many of the above mentioned configurations. It had only a marginal effect. Other DD variations included allowing for change of the structure in the wide region after the taper either along with optimization of the taper or exclusively. This did not lead to any functional devices either.

A mirrored structure which would allow for both the source and the objective to be placed in the narrow waveguide was also attempted. This could either be done by mirroring the taper right upon it reaching the maximum width or after allowing for propagation through some length of the wide waveguide. Unfortunately this tended to lead to a design where light was confined to the central part of the structure rather than expanded along the taper. Although never attempted it might have been worthwhile to include objectives not only in the narrow region but also in the wide central part of the structure. Although neither of these methods led to any functional structures, this last configuration became the basis of a different design configuration described below in section 7.3.1.

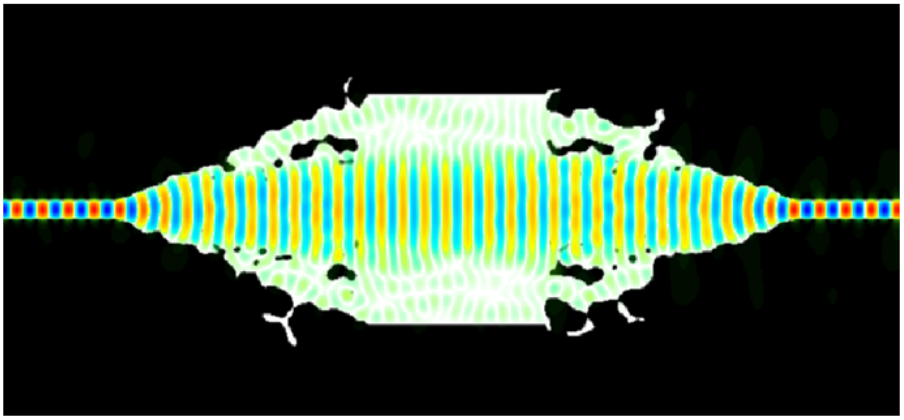


Figure 7.2.5: Illustrations of a mirrored taper structure constructed to allow for both the exciter and the objective to be placed in the narrow waveguide. Optimized design overlaid with the 2D FDTD calculated H_z field.

7.3 Small difference taper

To verify the postulate that it is the inability to excite the TE_0 mode in the wide waveguide that inhibits the TO from optimizing the compact taper structure, a number of smaller scale tests are made. Another benefit of working on the small scale is that all optimizations can be done in 3D without the need of any preliminary 2D optimization.

The first small scale example is to optimize a small taper between waveguides that are so narrow that the TE_0 can be excited in both of the photonic wires. Figure 7.3.1 shows the starting point structure for the optimization along with the resulting structure and the field simulation through both optimized and unoptimized device.

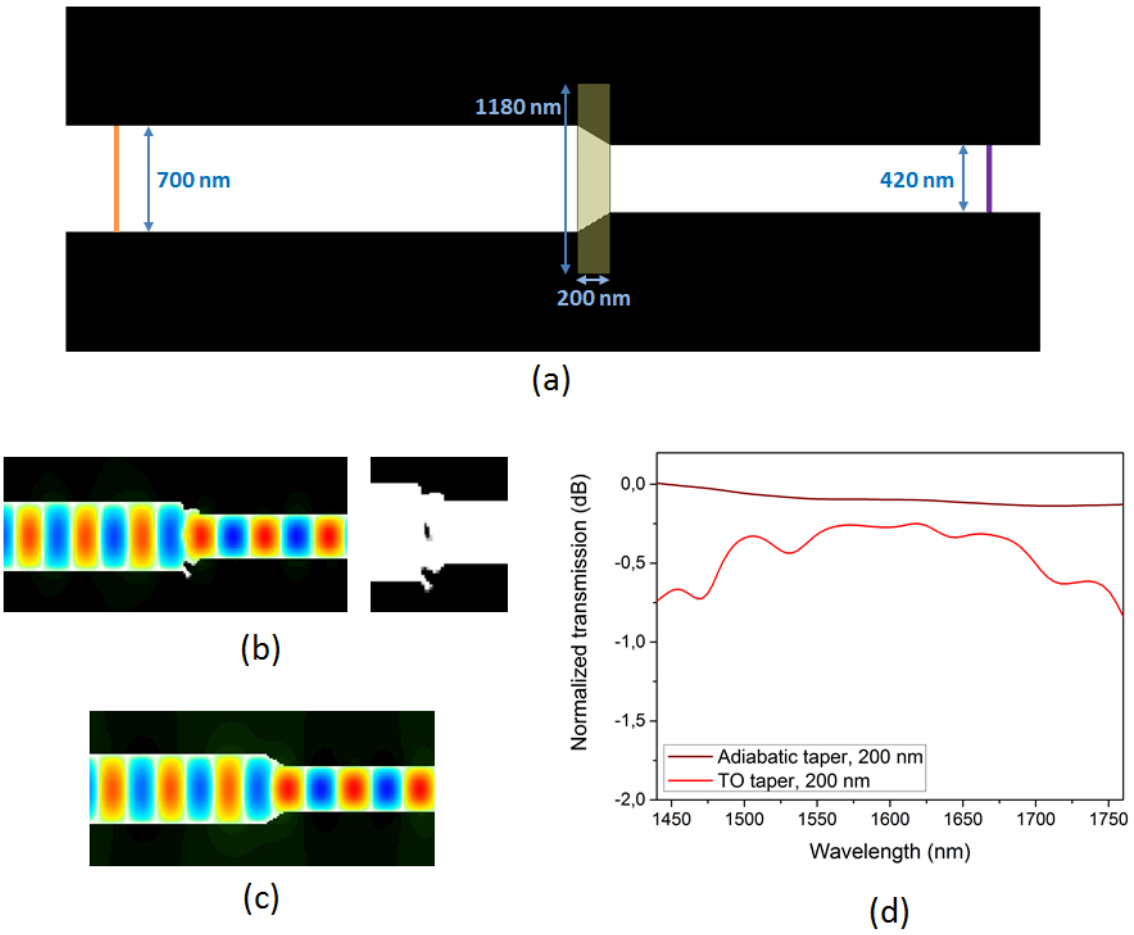


Figure 7.3.1: (a) Starting point structure for the optimization of a compact taper between two straight photonic wire waveguides carrying the TE_0 mode. (b) H_z field simulation through the optimized taper structure along with a close-up of the optimized region. (c) H_z field through the original adiabatic taper structure. (d) Normalized calculated 3D FDTD spectrum for transmission through the adiabatic and topology optimized tapers.

Optimizing from this much simpler starting point verifies that Phazor is indeed capable of handling this type of problem formulation, as long as the fundamental mode can be excited. Unfortunately this demand requires that the difference between the two straight waveguides is so small that even an untapered transition between the two will maintain the fundamental mode. In the case of the very functional adiabatic taper, TO will not be able to improve the performance as is evident from figure 7.3.1(d).

7.3.1 Stepwise taper design

A different approach, very similar to that presented in figure 7.1.2 is then attempted as a basis for a step-wise optimization. Although shown not to work on the large scale, the use of the adiabatic taper leading to a straight region of a wider TE_0 is useful for waveguides with little width difference. The first step of the set-up is shown in figure 7.3.2(a). The small taper region is optimized and subsequently the field-test is done by reimporting to a structure without the leading taper region, see figure 7.3.2(b).

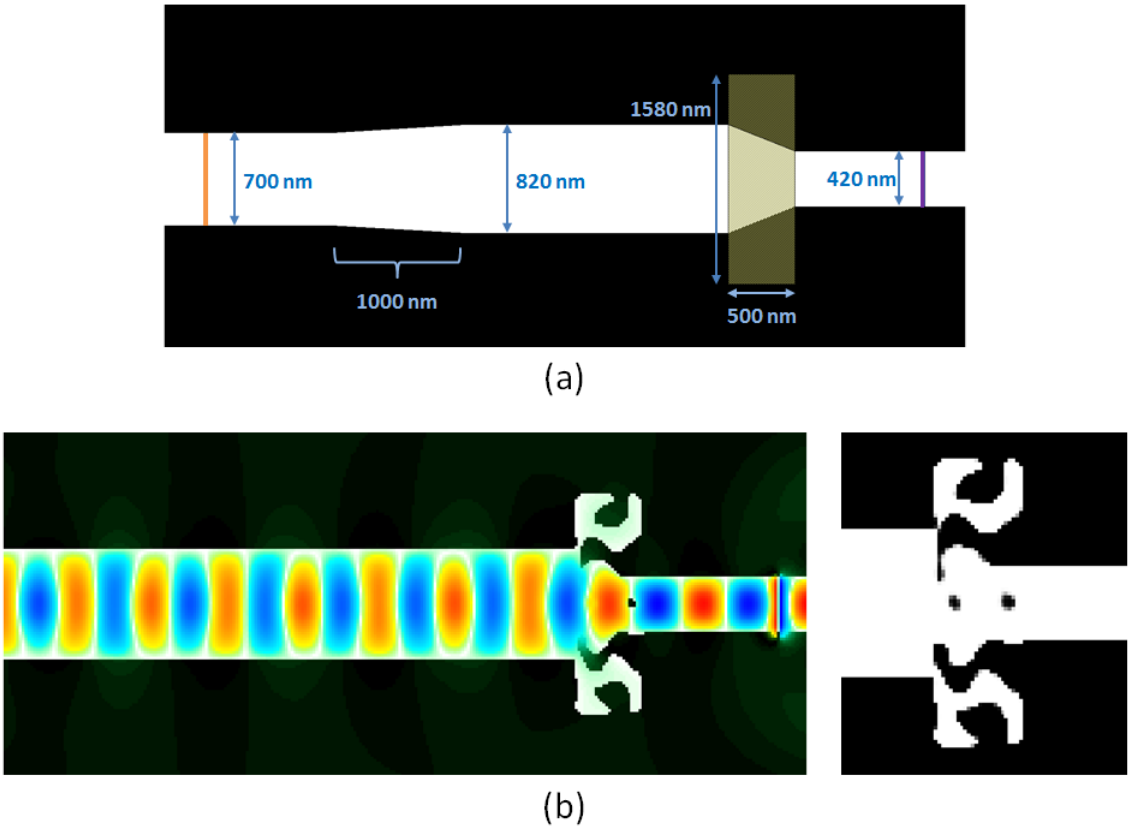


Figure 7.3.2: (a) Starting point structure for the first step of the stepwise optimization. By gently tapering the input side (left) from a width where the TE_0 mode can be successfully excited (orange) to a wider region, a bigger waveguide dimension difference is obtained. The design domain (yellow) is then placed around a shorter taper bringing the signal to a single mode photonic wire wherein the objective is places (purple). (b) Subsequent 3D FDTD H_z field simulation in a structure with only the optimized taper between straight waveguide regions. Presented along with a close-up of the optimized region.

Having optimized one small taper between two photonic wires very similar in width, it can now be used as the lead for the next structure. Beginning with the optimized region and then including an adiabatic taper, before the small region of optimization, a bigger width difference taper can be optimized. The second step optimization stating point and resulting design are shown in figure 7.3.3.

Potentially this stepwise optimization could be continued to an arbitrary width difference. It is however neither an elegant nor a practical solution, as the stepwise optimization is cumbersome for creating multiple starting points and time consuming for the many subsequent optimizations needed. Thus this approach serves only the purpose of demonstrating that it is possible to use TO to create low-loss compact tapers if the necessary starting point can be created. Figure 7.3.4 compares the stepwise optimized taper to an unoptimized taper of the same length and presents

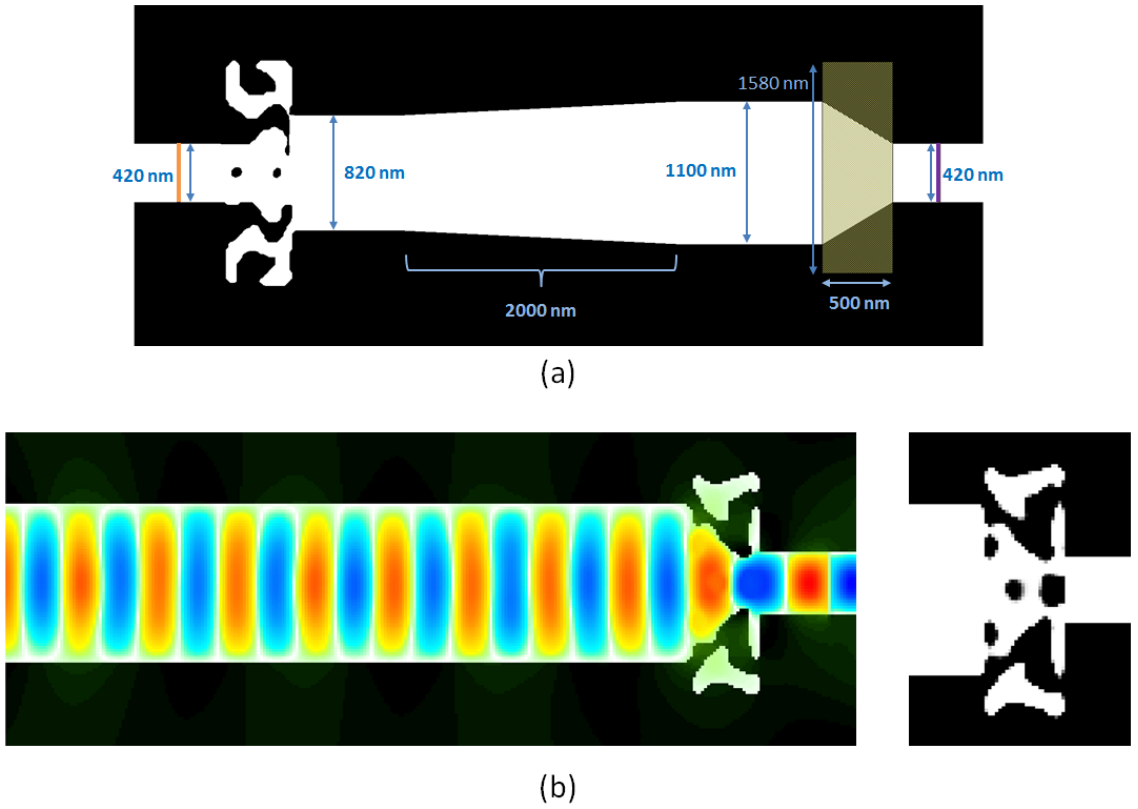


Figure 7.3.3: (a) Starting point structure for the second step of the stepwise optimization. The TE_0 mode is excited in the narrow waveguide (orange) and led to a wider waveguide using the optimized taper presented in figure 7.3.2. The mode is then gently tapered to a yet wider waveguide region before it is introduced to the short taper surrounded by the design domain (yellow). (b) 3D FDTD simulation of the H_z field of the optimized structure obtained from (a) but now connecting only straight photonic wires. Presented along with close-up of the optimized region.

the spectrum for the former case. It is evident that TO can be used to optimize such a structure given the possibility of exciting the TE_0. The spectrum verifies the broadband functionality, however some loss is present. Presumably some of this loss is a consequence of the stepwise optimization where the loss of any unideal features in an early step will be carried over into all subsequent steps of the optimization. This is an additional reason that it is not a practical approach. Furthermore fine-tuning the device in each steps is a time-consuming process which has not been prioritized; it is however likely that slightly better performance could be obtained even using the stepwise approach.

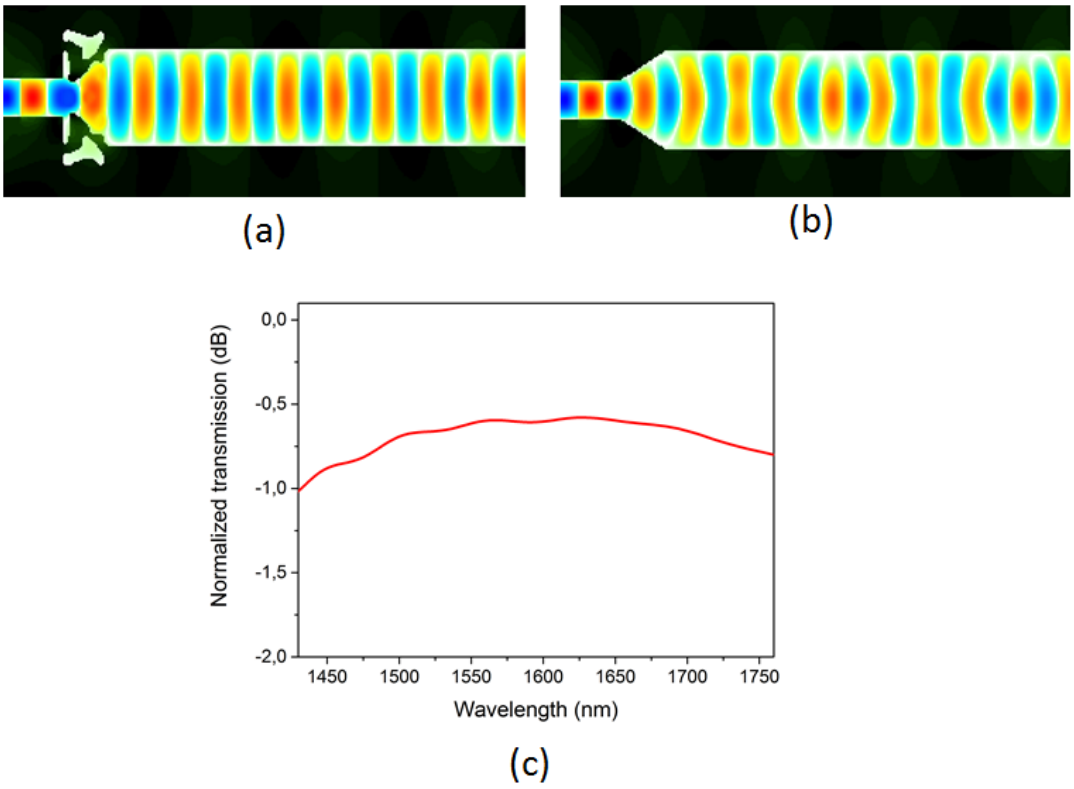


Figure 7.3.4: (a) Topology optimized 500 *nm taper between waveguides of width* 420 *nm and* 1100 *nm obtained using the stepwise optimization approach. Overlaid with 3D FDTD calculated H_z field. (b) H_z field through a* 500 *nm unoptimized taper between waveguides with the same dimensions as for the optimized case. (c) Transmission spectrum for the optimized taper normalized to that of a straight* 420 *nm photonic wire.*

7.4 Summary and discussion

Compact tapers are essential devices for the area-efficient packaging of PICs. Here it has been demonstrated, on a small scale, that TO can be used to create efficient transitions shorter than necessary for linear adiabatic tapering. There is no inherent limit to how large a waveguide width difference TO can be used to design transitions for. The current implementation of the TO method does, however, face the challenge of not having a mode exciter or objective. This means that the limiting factor is the ability to excite the TE_0 mode in a wide waveguide. Various approaches were attempted including a leading linear taper and manually exciting the mode profile by modifying the amplitude in every single point. The latter method worked in 2D but was not possible to scale to 3D. Topology optimization is expected to be a good tool for designing compact tapers and further investigation on overcoming the challenges of the current implementation is deemed to be worthwhile.

CHAPTER 8

CONTINUOUS CLADDING MODULATED BRAGG GRATINGS

8.1 Introduction

This chapter covers the part of the project that took place at the IBM T. J. Watson Research Center in Yorktown, New York. Although still in the field of integrated silicon photonics it does not rely on topology optimization and as such this chapter will include a bit more background material than the previous chapters containing results.

The devices under investigation are distributed Bragg reflectors (DBRs), or Bragg gratings, which are fundamental components for a number of optical devices. Distributed Bragg reflectors are optical wavelength filters which are created by periodically modulating the effective refractive index [66],[67]. This modulation can be obtained either by a variation of the material or the geometry of the waveguide, an example of the latter case is the 1D PhC where the photonic bandgap creates the reflection. At each interface between two different refractive indices a partial reflection occurs. At the Bragg wavelength, the reflections of the layers interfere constructively and form a strong reflector. In this manner the DBR creates a stopband in the spectrum of light incident on the grating. Only light of the Bragg wavelength is reflected and the rest of the spectrum is minimally affected by the grating, the concept is sketched in figure 8.1.1. The (DBR) allows for tuning of the bandwidth of the stop band by tuning of the amplitude of the index modulation and the Bragg wavelength by altering the period of the modulation.

Bragg gratings have been widely used in fibres [66][68] and also realized on chips. The uses for these reflectors are many such as strain sensing, partially reflecting mirrors, notch filters and in fact for add-, drop- wavelength multiplexing. The reasons for going on-chip are, as stated in chapter 1, in short that it allows for compactness and the integration with microelectronics. Several ways exist of creating DBRs in waveguides among which are stacks of intermittent materials [69], corrugation [70] and sidewall-modulated distributed resonant structures (S-DRS) [71],[72]. The S-DRS are of special interest to this project. They exploit the fact that changing the width of the waveguide also changes the effective refractive index, thus by periodically altering the width, making teeth on the sidewall, a Bragg grating is created. The concept is sketched in figure 8.1.2(b) as it would be designed. During the process of fabrication such sharp features will, however, be softened creating rounded corners. This discrepancy between the design and the fabricated structures

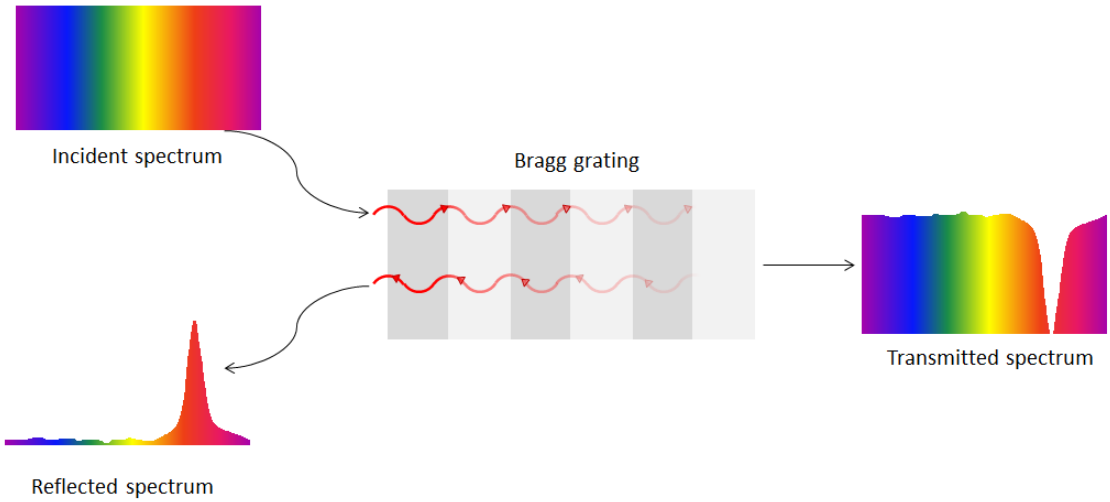


Figure 8.1.1: Sketches of the concept of a distributed Bragg reflector. Periodic variances in the effective refractive index leads to reflections at each interface. Due to constructive interference this reflection becomes strong and causes a stopband in the transmission spectrum at the Bragg wavelength.

causes challenges for the control of the Bragg grating functionality. However, one of the main benefits of the S-DRS is that the processing steps are few and relatively easy. During fabrication only a single lithography step is needed as only a single waveguide material needs to be present. Meanwhile there are limitations in this fabrication procedure. The coupling coefficient is determined by the amplitude of the sidewall modulation which is limited by resolution of the used lithography equipment.

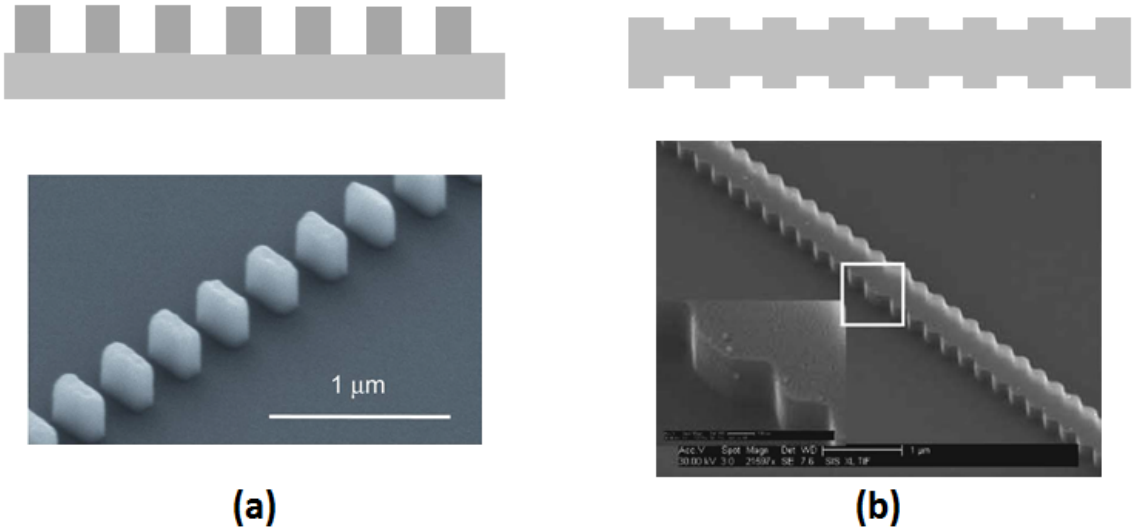


Figure 8.1.2: (a) On top a side-view sketch of the concept of a rib waveguide. A vertical Bragg grating functioning by alternating the materials of air and the guiding material which may or may not be the same a the substrate. Below an SEM image of such a waveguide taken directly from reference [69]. (b) On top a top-view sketch of the concept of a side-wall modulated Bragg grating below which is an SEM image of such a structure originating from reference [71].

In [73] a different scheme is suggested which retains the easiness of the fabrication but without the limitations of the S-DRS. The structures which modulate the effective index is in this case placed in the cladding rather than connected to the waveguide. The concept of cladding-modulated distributed resonant structures (C-DRS) is illustrated in figure 8.1.3. In this case the evenly spaced cylinders are placed on both sides of the waveguide and, when close enough for the evanescent tails of the propagating field to couple to them, they modulate the effective refractive index periodically. The work of [73] found that the coupling coefficient as well as the bandwidth of the stopband decrease exponentially as the distance from waveguide to cylinder is increased.

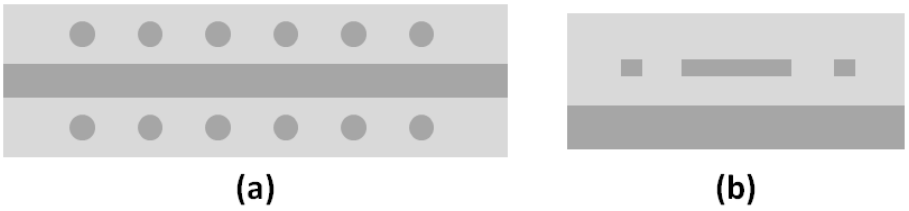


Figure 8.1.3: Illustration of the cladding modulated distributed resonant structure. Dark grey is silicon and light grey silica. (a) Top view of the structure showing the concept. (b) Side view illustrating the complete silica cladding.

The goal of the work presented here is to build upon the concept of C-DRS to make a type of Bragg gratings that are even easier to fabricate. This is intended by making the cladding modulators continuous rather than individual cylinders. In this manner the overall features are not as small and the geometries are easier to (re-)produce reliably.

The devices under investigation in this chapter are continuous cladding modulated distributed Bragg gratings (C-DBRs). Two different design types were considered, namely a step-index modulation and a continuous width modulation as illustrated in figure 8.1.4. The modulation induced by the step-index grating is more akin to the structure of [73], but suffer from the drawback that the perpendicular angles of the design are not easily recreated during fabrication. This leads to a discrepancy between simulations and experimental results whereby the design task becomes more difficult. It is the hypothesis that these issues are alleviated by the continuous width grating, however the more gradual index change might lead to weaker reflections and a more complex relation between reflection and design parameters.

Simulations were performed using the commercial tool Lumerical Mode Solutions employing its eigenmode expansion (EME) solver. In broad terms the method is a fully vectorial, bi-directional tool for the solving of Maxwell's equations. It functions by dividing the structure into cells in which the refractive index is constant in the propagation direction and then solves for the modes at the interfaces. In this manner a basis set of eignemodes is formed with the main benefit being that increasing the number of periods require very little additional computation making it ideal for periodic structures such as gratings [74]. The devices are defined from simple geometrical blocks along with a specification of the periodic structure cell and an indication of how many times it should be repeated. Figure 8.1.6 presents the two types of structures processed with Mode Solutions. These simulations directly return the calculated transmission spectrum of the device. To get an understanding of the functionality and behaviour of these structures, a number of simulations were made mapping out the effect of changing the various parameters.

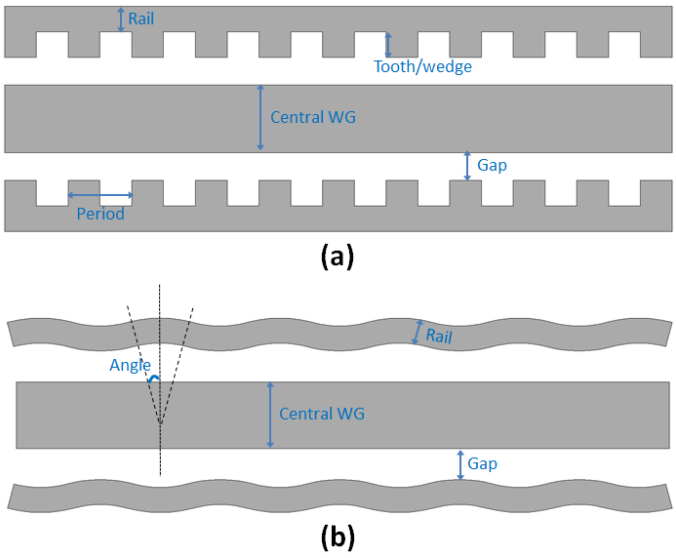


Figure 8.1.4: Illustrations of the two types of cladding modulated continuous grating devices. (a) The step index grating. (b) The continuous width grating.

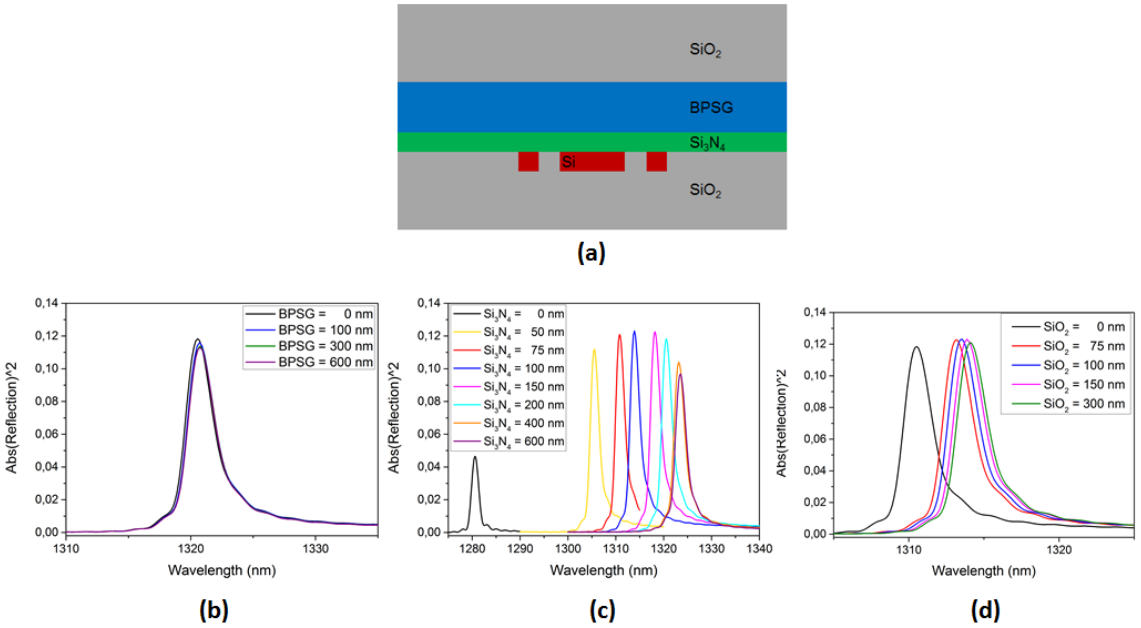


Figure 8.1.5: (a) Cross-section view of the cladding modulated Bragg grating showing the various coatings. (b) Spectrum of the grating structure for differing thickness of the BPSG layer while $t_{Si_3N_4} = 200\,nm$ and $t_{SiO_2} = 150\,nm$. (c) Altering Si_3N_4 thickness while the other layers are $t_{BPSG} = 0\,nm$ and $t_{SiO_2} = 150\,nm$. (d) Changing thickness of the SiO_2 coating while the other two layers have fixed heights of $t_{Si_3N_4} = 100\,nm$ and $t_{BPSG} = 0\,nm$.

The devices are fabricated on an SOI substrate, with SiO_2 also filling the area between the waveguides and reaching the same height as the Si. Different coating layers of Si_3N_4, BPSG and SiO_2 have been investigated. A schematic cross-section with a three layer coating is shown in figure 8.1.5 along with simulations of the effect of altering the thickness of the various layers. It is found that the BPSG layer has very little effect even for the extremes of no layer or a very thick coating. Varying the Si_3N_4 layer alters the resonance frequency, but does not change the reflection strength significantly unless it is removed completely; too thick a layer does however seem to dampen the effect. The effects are similar for the SiO_2 layer although even smaller. Altering the thickness causes only little difference as long as the coating is present.

In the following sections the two device designs, the stepwise and the continuous index change, will be discussed separately. In both cases the results of the simulations will be presented followed by the experimental results from the fabricated devices. For specific subsets of the measured data a direct comparison between simulations and experimental results are included. Finally the findings will be summarized in section 8.4 looking at the performance of the two device types and correspondence between simulations and experiments.

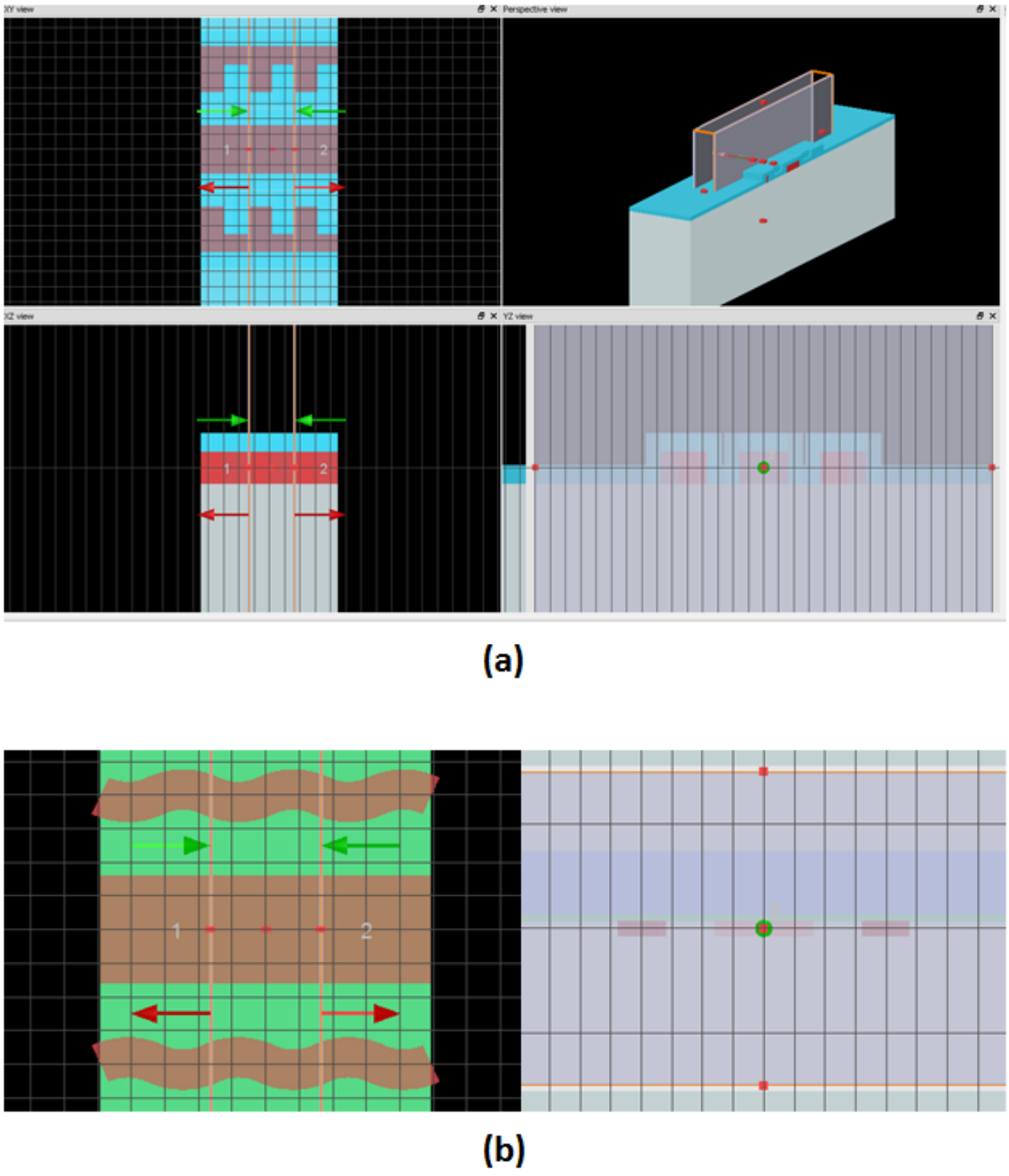


Figure 8.1.6: Screen dumps from Lumerical Mode Solutions showing how the two types of structures have been defined for calculation. (a) The step index grating including all four structural viewpoints. Here the structure is shown with a 100 *nm uniform coating. (b) The continuous width grating for only top view and end view of the structure here presented with all three coating layers. The orange lines on the left row of pictures indicate the cell which is being repeated. The arrows indicates the calculations of the transmitted and reflected light which will be performed in each direction at the interfaces of the cells.*

8.2 Step index gratings

A previous iteration of the C-DBR had been attempted at IBM. It did indicate that the endeavour would be possible, meanwhile design errors led to the effect being minuscule. Hence the investigation undertaken was twofold. It aimed to determine if the existing fabrication mask could be used with minor alterations to the fabrication procedure, as well as to see how a stronger effect could be gained for a new mask. Because of the already existing work, a thorough investigation of the effect of parameter change was not the goal of the work with the step index C-DBR. During the simulations a number of trends were, however, observed and they will be presented here prior to the experimental results.

8.2.1 Step index simulations

Initially there were two different types of step index gratings one with the teeth facing inwards as shown in figure 8.1.4(a) and one where the teeth face outwards being placed on the opposite side of the rail. Figure 8.2.1 offers a comparison of simulated reflection spectra of the two grating configurations. All parameters but the length are the same. Neither grating exhibits a strong reflection but notably the reflection of the outwards grating is 5 times weaker in spite of it being 25 times longer than the inwards facing C-DBR. Hence the design with outward facing teeth was thus not perused further in this project.

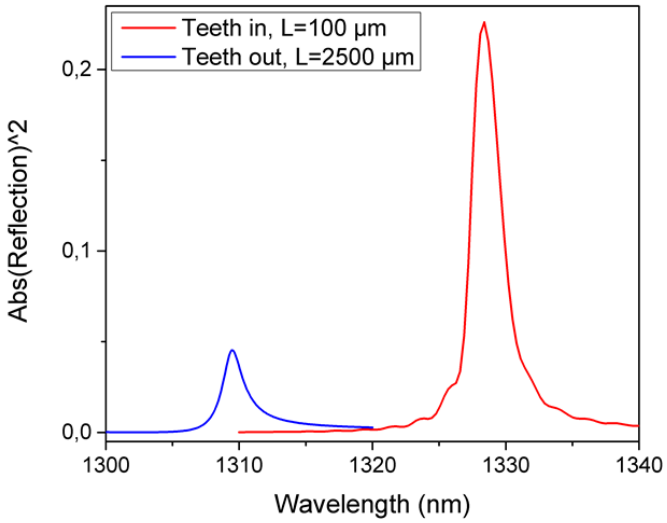


Figure 8.2.1: Simulated reflection spectra for two step index continuous cladding modulated gratings with the teeth facing in and out, respectively. Aside from the lengths of the gratings all other parameters were the same: $WG = 350\,nm$, $Gap = 150\,nm$, $P = 300\,nm$, $wedge = 150\,nm$, and $Rail = 120\,nm$.

In all of these designs the rail width was kept at 120 nm. This was done because the purpose of this rail is solely to render the structure continuous and thereby easier to fabricate. 120 nm was the smallest feature size that could safely be assumed to be producible and making it larger was deemed unlikely to reveal better results. Changing the size of the wedges, or teeth, on the rail was however a tunable parameter. Figure 8.2.2 compares the reflection spectra for different wedge depths for various design configurations. For all of the datasets, in spite of their great variance in strength, a trend appears, that the larger teeth lead to a weaker reflection. Figure 8.2.2(a) does however indicate that there may be an optimum as the reflection is slightly weaker for the 100 nm wedges than 110 nm, this effect is however minor.

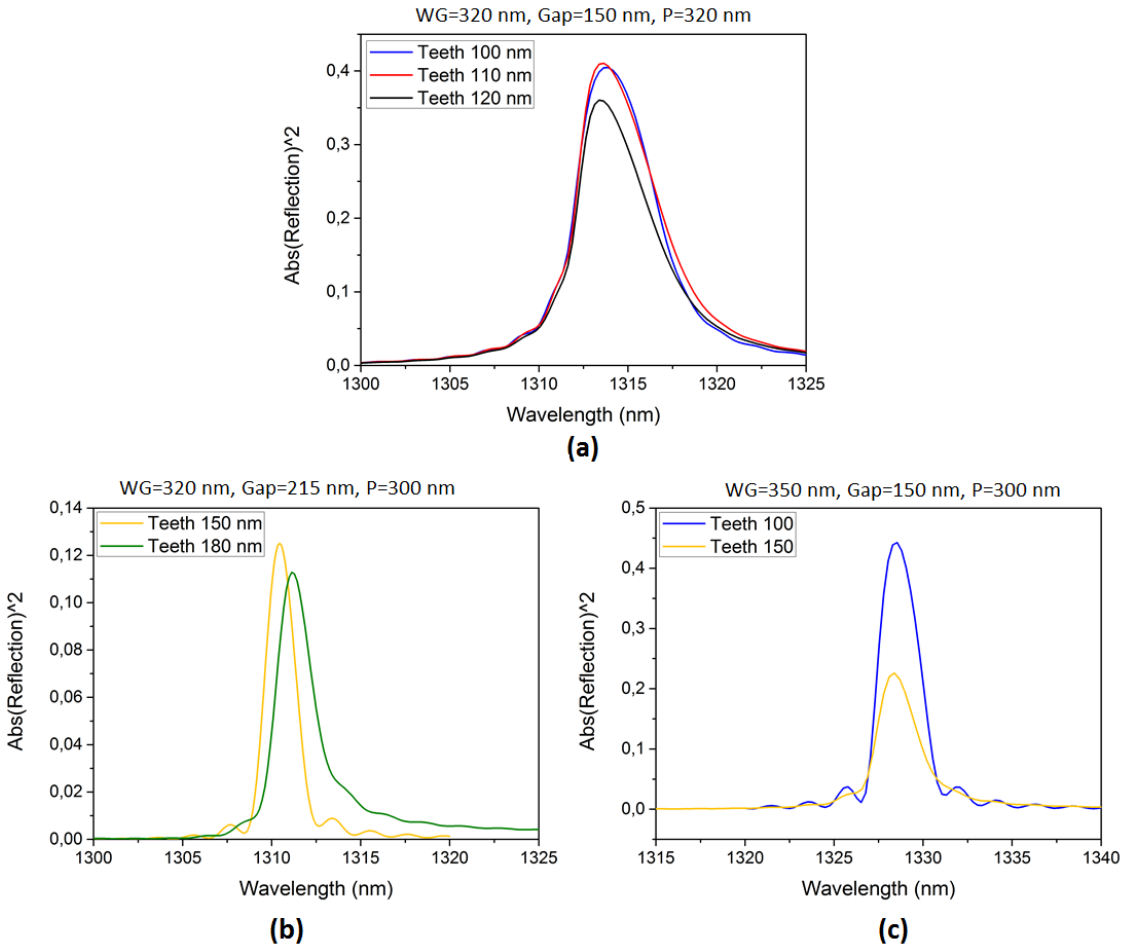


Figure 8.2.2: Simulations of the reflection from the continuous step index gratings when the tooth size, that is the depth of the wedge, is altered. Within each graph all other parameters than the wedge depth are fixed and the values are specified above. Gap width, width of the center waveguide and the period have been altered between subfigure (a), (b) and (c).

Two other parameters were investigated in some depth, namely the period of the grating and the width of the central waveguide. Both of the parameters influence the reflection wavelength and strength. The central waveguide width will in practice be limited to ensure that higher order modes will not be present, however understanding how it influence the performance is still valuable. Figure 8.2.3 shows how the reflections develop as the period is changed. This relation was investigated for two different center waveguide widths of $WG = 320\,\text{nm}$ and $WG = 360\,\text{nm}$, all other parameters are fixed. Although the development is not identical for the two waveguide widths, the trends are similar. The reflection wavelength increases linearly with the period. While strength and bandwidth do increase monotonously, their effects are not as linear and especially for high wavelength the increase seems to stagnate. These tendencies are underlined in figure 8.2.4(a) where the waveguide width is the focus. Here too it is clear that increasing the period will increase the reflection wavelength linearly. At the same time it is found that a change in the width of the central waveguide will lead to a shift in reflection wavelength. Namely narrowing the waveguide shifts the reflection to lower wavelength. The magnitude of the downward shift is equal for all periods. Figure 8.2.4(b) focuses on the width of the central waveguide alone and it is apparent that the transmission wavelength increases as the width is increased, however this also leads to a loss of reflection strength and bandwidth.

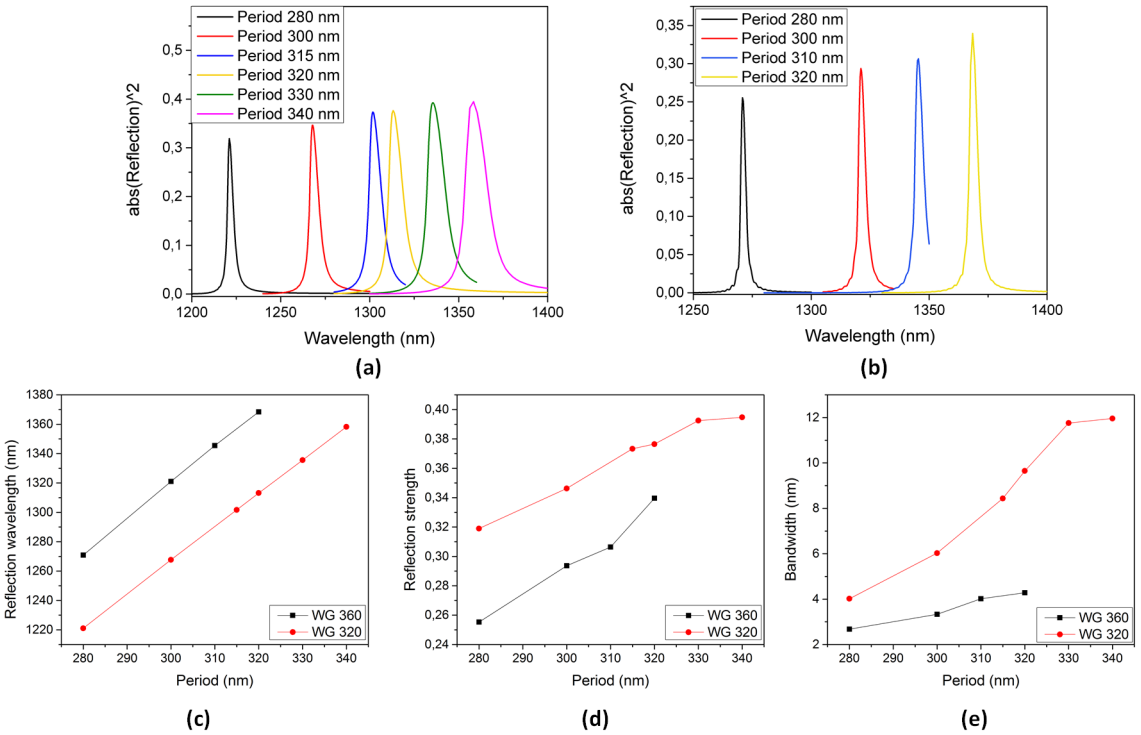


Figure 8.2.3: Simulations of the reflection spectra of step index gratings with the period varied while all other parameters are fixed, for (a) a center waveguide width of $WG = 320\,nm$ while for (b) it is $WG = 360\,nm$. (b-d) show the development of the reflection wavelength, strength and bandwidth, respectively, as a function of the grating period.

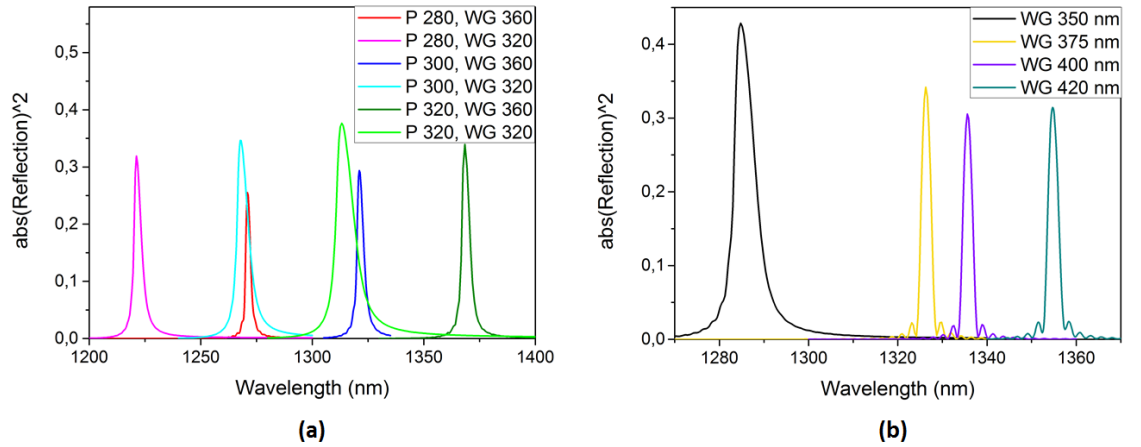


Figure 8.2.4: (a) Simulated reflection spectra for step index cladding modulated devices under change of period and center waveguide width, while other parameters are kept fixed. Light and dark colours have the same waveguide width respectively. (b) Development of the reflection spectra as the width of the central waveguide is varied. $P = 300\,nm$, $Gap = 150\,nm$, and $wedge = 150\,nm$.

8.2.2 Step index measurements

The gratings were fabricated in the IBM T. J. Watson research centre, without my personal involvement in the processing, an SEM image of the resulting structure is shown in figure 8.2.5. It is clear that the edges are softened and corners less well-defined than in the mask design. This discrepancy is expected to lead to a loss of correspondence between experimental and theoretical data. The previous analysis is still of relevance, though, both in choosing the parameters for the device design and because the trends for parameter changes are likely to exhibit similarities to the simulations even if the spectra change.

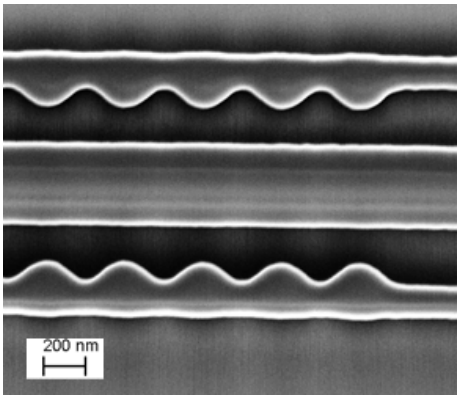


Figure 8.2.5: SEM image of a fabricated step index continuous cladding modulated distributed Bragg reflector after etching but before coating. Fabrication was successful yet the corners of the structures are softened compared to the mask design sketched in figure 8.1.4 and 8.1.6.

The fabricated step index C-DBRs all have a central waveguide width of 320 nm, rail width of 120 nm and wedge depth of 100 nm. In the first part of the analysis presented here all structures have a length of $L = 600\,\mu$m while the period and the gap are altered.

Several of the spectra exhibit double reflections at spaced wavelengths as exemplified in figure 8.2.6(a). These additional reflections may be either second order Bragg reflection lines or an indication that the signal has coupled to other modes. These additional reflections did not appear in any of the simulations. The first reflection does occurs consistently in the measured spectra at a wavelength near 1255 nm. It was suspected to be an artefact of the set-up and this was confirmed through reference measurements. Figure 8.2.6(c) and (d) show the reflection position and strength as a function of the period, respectively; the before-mentioned double reflections have been excluded. Aside from a few outliers, the reflection wavelength increases monotonically with the period as was expected from the simulations. As is made clear by the direct comparison to the simulations in figure 8.2.6(d), the development in reflection position is, however, not entirely linear. In the central period region the dependence is linear with a rate similar to those found in the simulations, however for both lower and higher periods discrepancies are found. Furthermore all positions seem to be shifted towards higher wavelengths. This is likely to be due to changes in the structures introduced during the fabrication. The overall shift is presumably due to a consistent change in the dimensions - as the shift is upwards in period it is likely that it is cause by line-broadening during the etching or a change in the coating layer thickness. The lack of linearity is likely to be due to the rounding of the structures, and it may be speculated that the effect is most severe for the shorter periods where structures are more closely spaced. In the experimental re-

sults another discontinuity is also observed where the reflection moves back to lower wavelengths for periods increased beyond 340 nm. These wavelengths were not originally examined during simulations as they were expected to fall outside the possible measurement range. The larger periods were, however, included on the mask to account for potential inconsistency between simulated and fabricated results. These large period reflections exhibit the greatest reflection strength by far and do all exhibit double reflections as well, see figure 8.2.6(b). Due to the inconsistency of these results they were compared across several chips and the same behaviour was exhibited on all of them. Additional simulations were made in the search of these reflection at lower wavelengths, but they do not appear; as previously speculated it may be a matter of coupling to a higher order mode. For periods below 360 nm the strength of the reflection exhibit no dependence on the period.

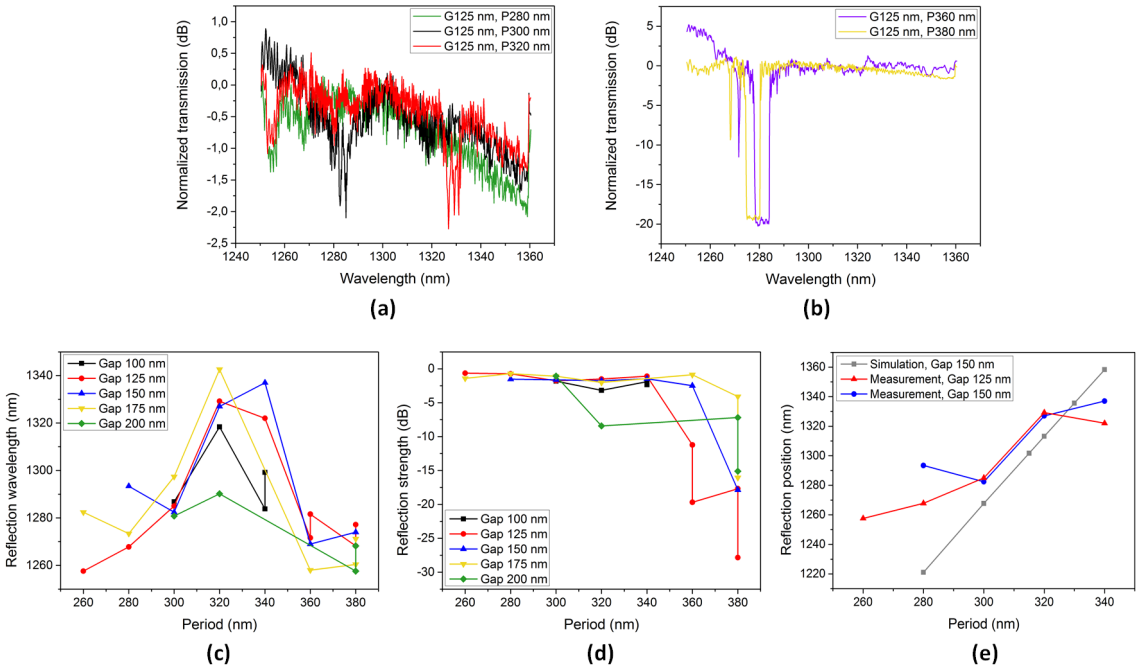


Figure 8.2.6: (a,b) Example spectra for measurements of step index modulated reflection gratings recorded for various periods. All other parameters are fixed: $WG = 320\,nm$, $Rail = 120\,nm$, $wedge = 100\,nm$, and $Gap = 125\,nm$. (c) The wavelength of the reflection and (c) the reflection strength as a function of the period. The various datasets show the dependence for differing gap sizes. (e) Shows a direct comparison between measured and simulated results.

Figure 8.2.7 shows the correlation between the reflections and the size of the gap between the central waveguide and the cladding structures. Although there are slight deviances, the overall trend is that the gap does not influence neither the position nor the strength of the reflection. It is the hypothesis that for the periods where this is not the case, differences are due to the low resolution and the noise in the measurement set-up as well as imperfections introduced during fabrication. Due to time constraints it was unfortunately not possible to measure on enough samples to collect the statistics needed to verify or dismiss this assumption.

The effect of the length of the grating was cursorily studied. From theory and simulation, it is expected that the longer gratings will exhibit stronger reflections and a narrower bandwidth. A more thorough numerical study of the length is conducted for the continuous width C-DBRs and presented in the next section, see figure 8.3.1. Only two lengths were investigated experimentally, namely $L = 100\,\mu m$ and $L = 600\,\mu m$; figure 8.2.8 presents a comparison of their performance. All

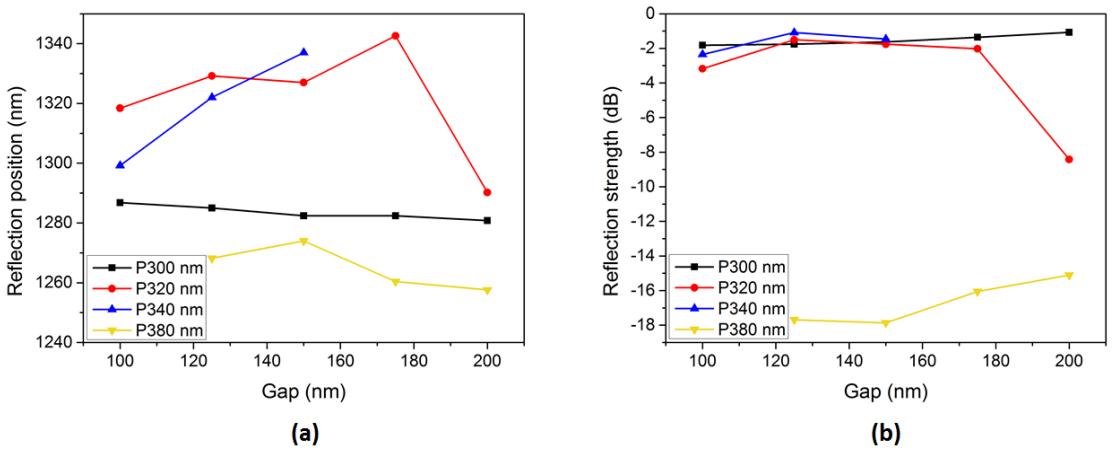


Figure 8.2.7: Dependence of (a) the reflection position and (b) reflection strength as a function of the width of the gap for a step index continuous cladding modulated distributed Bragg reflector. Other design parameters are fixed to $WG = 320\,nm$, $Rail = 120\,nm$, and $wedge = 100\,nm$, while different datasets are shown for changes of the period.

simulations were conducted at a length of $L = 100\,\text{nm}$ and this was deemed to be plenty for generating strong reflections. Meanwhile most of the spectra exhibited only very weak reflectivity for the $L = 600\,\text{nm}$ case and have completely disappeared for the shorter gratings.

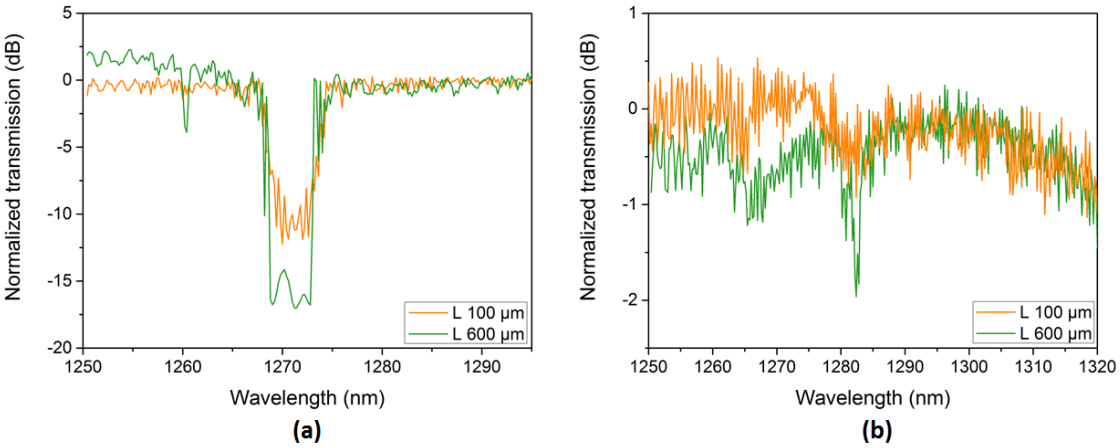


Figure 8.2.8: Comparison of spectra recorded for the same type of step index gratings for two different grating lengths of $L = 100\,\mu m$ and $L = 600\,\mu m$, respectively. (a) The strongest grating case of $P = 380\,nm$ and $G = 175\,nm$. (b) For a weaker grating case of $P = 300\,nm$ and $G = 150\,nm$. All spectra are normalized to the value measured at $1300\,nm$, for that particular sweep, where the equipment was stabilized before the spectra were recorded.

The strongest recorded gratings, found for a period of $P = 380\,\text{nm}$ and shown in 8.2.8(a), illustrate the expected relation between different grating lengths. The longer grating exhibits a stronger reflection and a marginally smaller bandwidth. Figure 8.2.8(b) compares two weaker reflections, showing a much less clear relation. The reflection from the shorter grating is barely present, however it is noticeable that the bandwidth is larger. More thorough parametric studies were not possible for the shorter gratings because of the weak reflections being drowned by the high noise floor.

8.3 Continuous width grating

To the best of our knowledge the scheme of a continuous width C-DBR reflector grating has not previously been attempted. A somewhat more thorough investigation of the development of the reflection spectrum, when altering various parameters, was therefore performed. The results of the Lumerical ModeSolutions simulations are presented first and are followed by experimental results from the fabricated devices. Subsequently a few additional simulations were made so make direct comparisons to the experimental data, these results will be shown along the way.

8.3.1 Continuous width simulations

A number of parameters were investigated for the continuous width C-DBR system including length (L), rail width ($Rail$), Period (P), angle ($Angle$) and width of the gap, considered the shortest distance between the center waveguide and the rail occurring during a period, (Gap). No investigation of the central waveguide width was undertaken, rather a standardised value of $WG = 320\,\text{nm}$ was used as it was tested before at IBM and it is ensured to not include higher order modes in the on-chip set-up used. Furthermore the same period, $P = 329\,\text{nm}$, was utilized for all other investigations than that of the period itself. The length of $L = 100\,\mu\text{m}$ was also a standard for these simulations being changed only when it was the parameter under investigation itself.

Beginning with a test with straight forward results, figure 8.3.1 shows the reflection spectra of a continuous width C-DBR for numerous grating lengths. The reflection strength increases with the length of the grating until it stagnates for very long gratings, as might be expected. The bandwidth decreases to a limit of $\approx 5\,\text{nm}$.

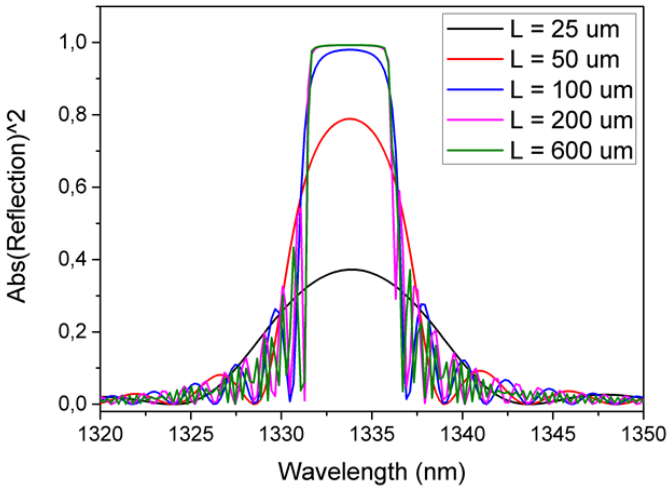


Figure 8.3.1: Reflection spectra of a continuous width cladding modulated distributed Bragg reflector. The length of the grating is varied while all other parameters are fixed at Angle = 15°, Rail = 120 nm and Gap = 150 nm.

Increasing the width of the rail leads to an increase in the reflection wavelength and the bandwidth, see figure 8.3.2. The reflection strength does however have an optimum exhibited near $Rail = 120\,\text{nm}$. It has not been thoroughly investigated, but there is a possibility that this optimum depends on the angle. If the angle is large, whereby the modulation on the rail becomes small, making the grating wider may quench the effect even more. For the case of the very narrow rail there will be less material for the evanescent modes of the center waveguide to couple too, if there is too little the effect will be diminished.

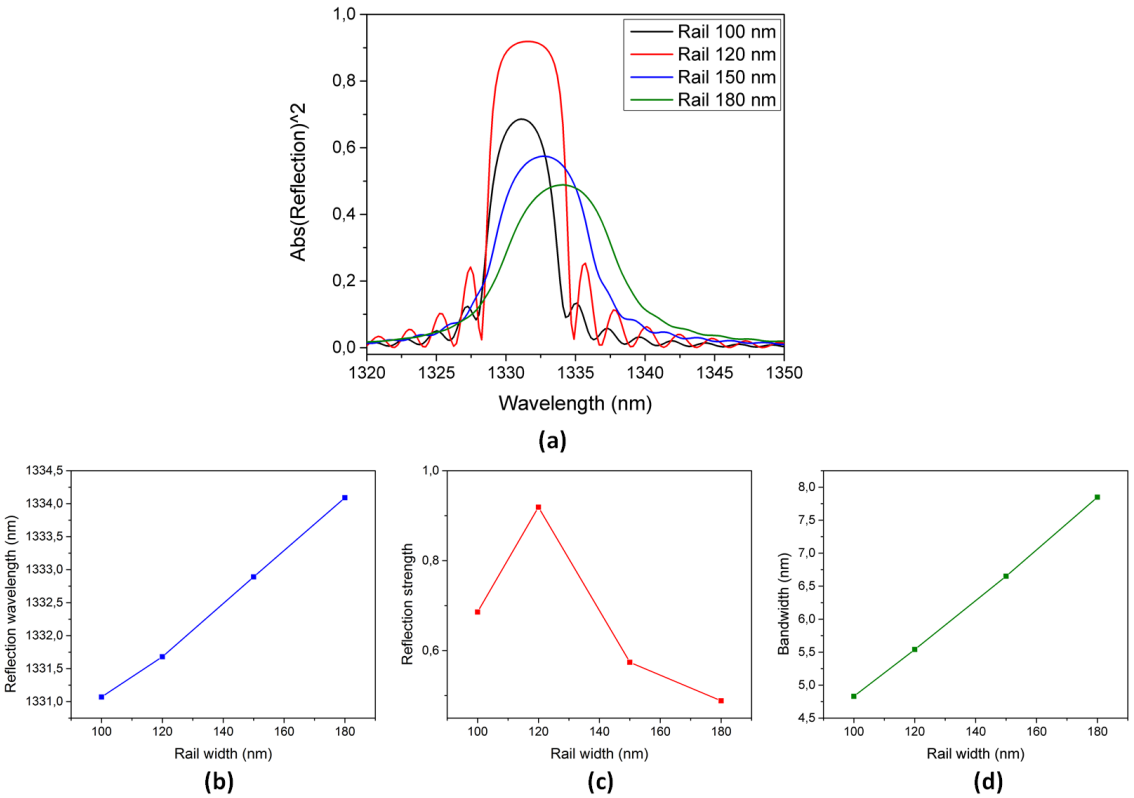


Figure 8.3.2: (a) Reflection spectra of the continuous width C-DBRs for various widths of the rail while other parameters are kept fixed at Angle = 45° and Gap = 170 nm. (b-d) show the development of reflection wavelength, reflection strength and bandwidth as a function of the rail width, respectively.

The period was investigated for a half-circle grating structure, that is to say that the angle is $Angle = 90°$. Unfortunately keeping this fixed and specifying the period of the structures the rail width was allowed to fluctuate. As for the case of the step index C-DBRs, the reflection wavelength is increased as the period is increased, as shown in figure 8.3.3. The reflection strength experiences an abrupt increase for the higher periods, but this also coincides with the wider rails. It is thus likely that the increase in strength comes from the rail width change because the design is closer to the optimum for this particular arc angle. Regardless it is concluded that the reflection wavelength can be controlled by altering the period, in correspondence with what was observed for the step-index grating.

Investigating the angle dependence of the grating reflection reveals that the device exhibits complex interactions between the various parameters. In figure 8.3.4(a) it is seen that, in regards to reflection strength, there is an optimum angle around 15°. Lowering the angle just by a few degrees does, however, lower the signal by $\sim 2/3$. More tolerance is exhibited for the increase in angle where the signal is lowered less drastically needing an increase of more than 70° for the same loss of the reflection strength. The former effect may readily be explained by the angles influence on the amplitude of the rail modulation. Very small angles will lead to the rail being practically a straight waveguide which would naturally induce no grating effect at all. It may be speculated that the loss for the higher angles originate from a slower change in the refractive index not leading to reflections. This hypothesis is however disputed by the results shown in figure 8.3.4(c) where it is found that when the gap is increased to $Gap = 190$ nm, the grating is strongest for a larger angle

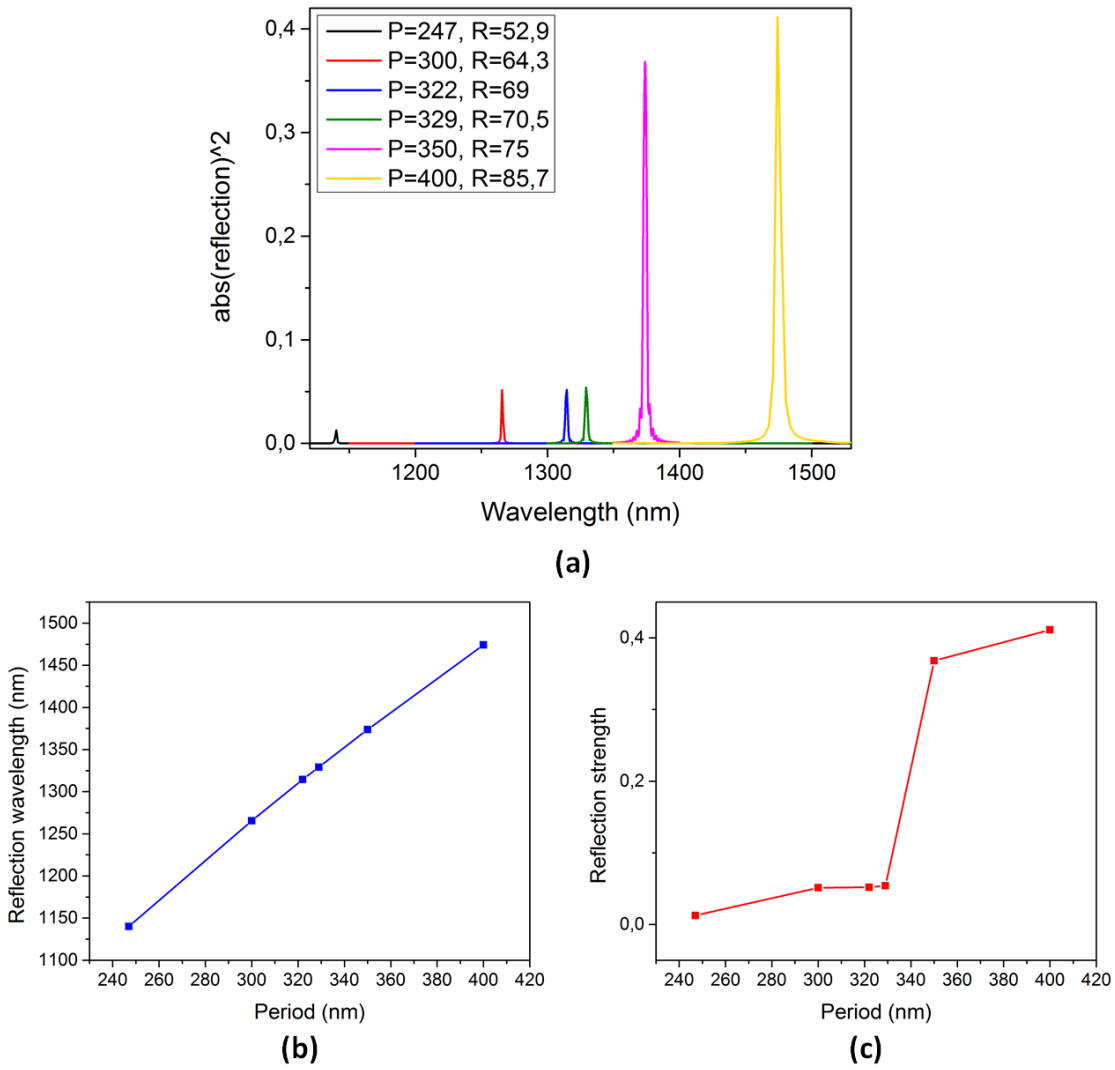


Figure 8.3.3: (a) Reflection spectra of the continuous width cladding modulated distributed Bragg reflector as the period is changed. Unfortunately it is not an isolated parameter investigation as also the rail width is varied throughout. Angle = 90° and Gap = 150 nm were however fixed parameters. (b,c) The development of reflection wavelength and strength as a function of the period. The rail width is not indicated in these plots although they are suspected to have noticeable influence on the reflection strength.

of *Angle* = 45°. It thus becomes apparent that although the trends are similar, see figure 8.3.4(d-f), the angle which leads to the strongest gratings and greatest bandwidth alike, shifts when the gap is altered.

The reverse mapping of the spectra of various gap differences for fixed angles are shown in figure 8.3.5. Trends are similar to those of the angles. The reflection wavelength decreases when the gap increases but for the strength and bandwidth there exists a parameter combination leading to a peak value. The change in the reflection strength is less abrupt than for the angle.

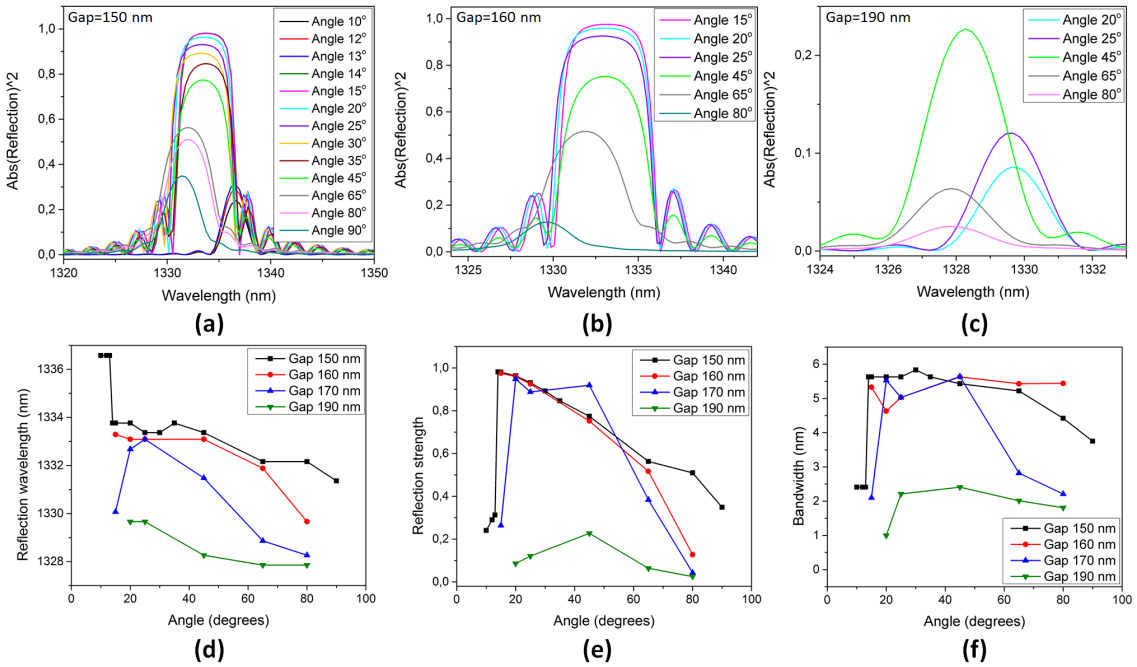


Figure 8.3.4: (a-c) Reflection spectra of the continuous width cladding modulated distributed Bragg reflector for various arc angles, portrayed at three different gap distances. (d-f) Development of the reflection wavelength, strength, and bandwidth as a function of the angle for four different gap distances.

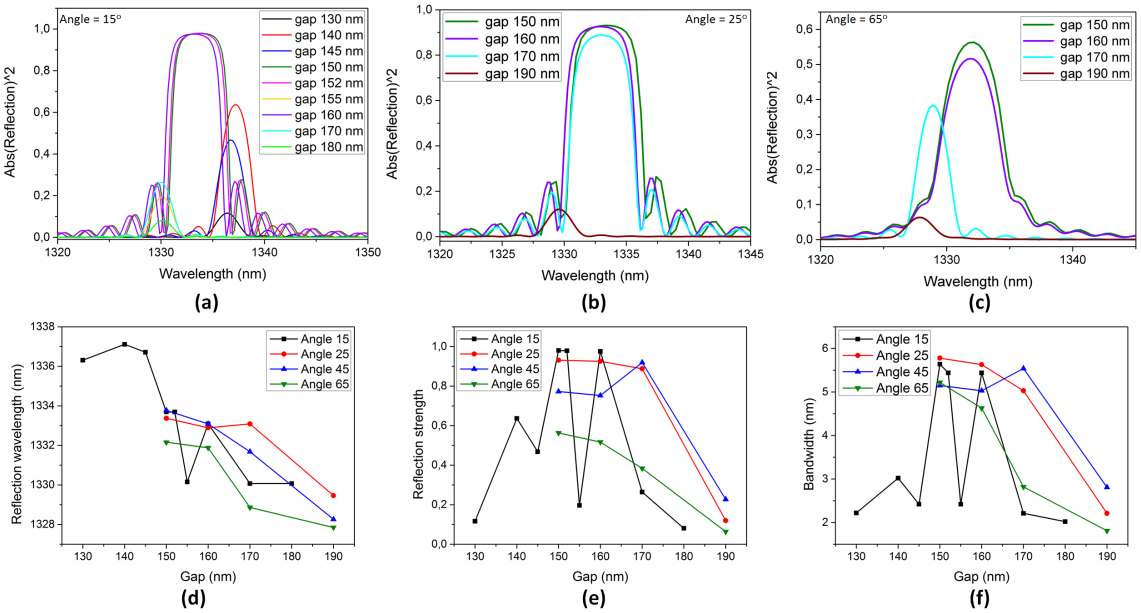


Figure 8.3.5: (a-c) Reflection spectra of the continuous width cladding modulated distributed Bragg reflectors when changing the gap width shown for three different grating angles. (d-f) Development of the reflection wavelength, strength, and bandwidth as a function of the gap width for four different structure angles.

This analysis of the various design parameters has shown that the relations are complex and that the designs will need to be carefully planned to get the desired functionality. However there is no doubt that this type of devices is functional and by fine-tuning the structures the reflection strength and wavelength can be engineered to fit a specific device.

8.3.2 Continuous width measurements

These structure, too, were fabricated at the clean room facilities of the IBM T. J. Watson research center, after my departure, and were subsequently characterized at DTU. An SEM image of the structure after etching and before cladding is shown in figure 8.3.6. It is apparent that the design has been transferred well to the pattern and in comparison to figure 8.2.5 the final structure more closely resembles the original design. In addition to the angle formulation of the continuous width C-DBR described above for the simulations, the 'waves' on the rails were also defined, on the mask, by a sine modulation. Results of both device types will be presented here.

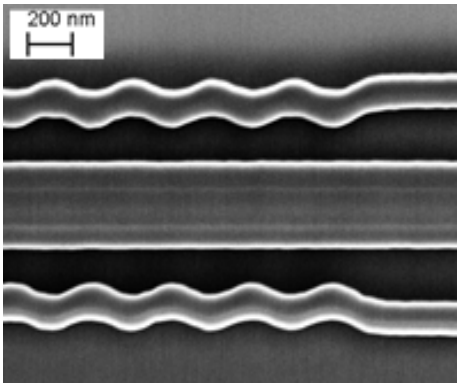


Figure 8.3.6: SEM image of the continuous width cladding modulated distributed resonant grating, the structures have been fabricated with great resemblance to the original designs as sketched in figure 8.1.4 and 8.1.6.

As the simulations showed a length of $L = 100\,\mu$m to be enough to give rise to strong reflections, most devices were fabricated for this parameter. However, as the measurements for the step-index gratings revealed greater losses than originally expected, the longer gratings with a length of $L = 600\,\mu$m are examined first.

Figure 8.3.7 presents example spectra along with graphs summing the development of the reflection position and strength as a function of the period. As expected from the simulations, the wavelength at which the reflection appears increase as the period becomes larger. A cut-off is seen for periods larger than 340 nm as the reflection moves outside the range of the laser. As the period is increased further another peak becomes visible at lower wavelengths, though. For the gap of 170 nm multiple dips occurred in the spectrum, however in all cases the recorded main reflection maintained a linear dependence on the period. Unlike for the case of the step index gratings, a tendency to deviate from the linearity when not at the central period was not exhibited. Figure 8.3.8(a) and (c) show the reflection positions as function of gap and angle respectively. A slight tendency towards shifting the reflection to lower wavelengths is observed for an increase of the gap. A similar, although weaker, relation is found for the angle. These relations are in correspondence with the simulation results presented in figures 8.3.5(d) and 8.3.4(d). The lower number of datapoints for the measured data may be masking any deviations that were found in the simulations, but the overall correspondence between the trends is evident.

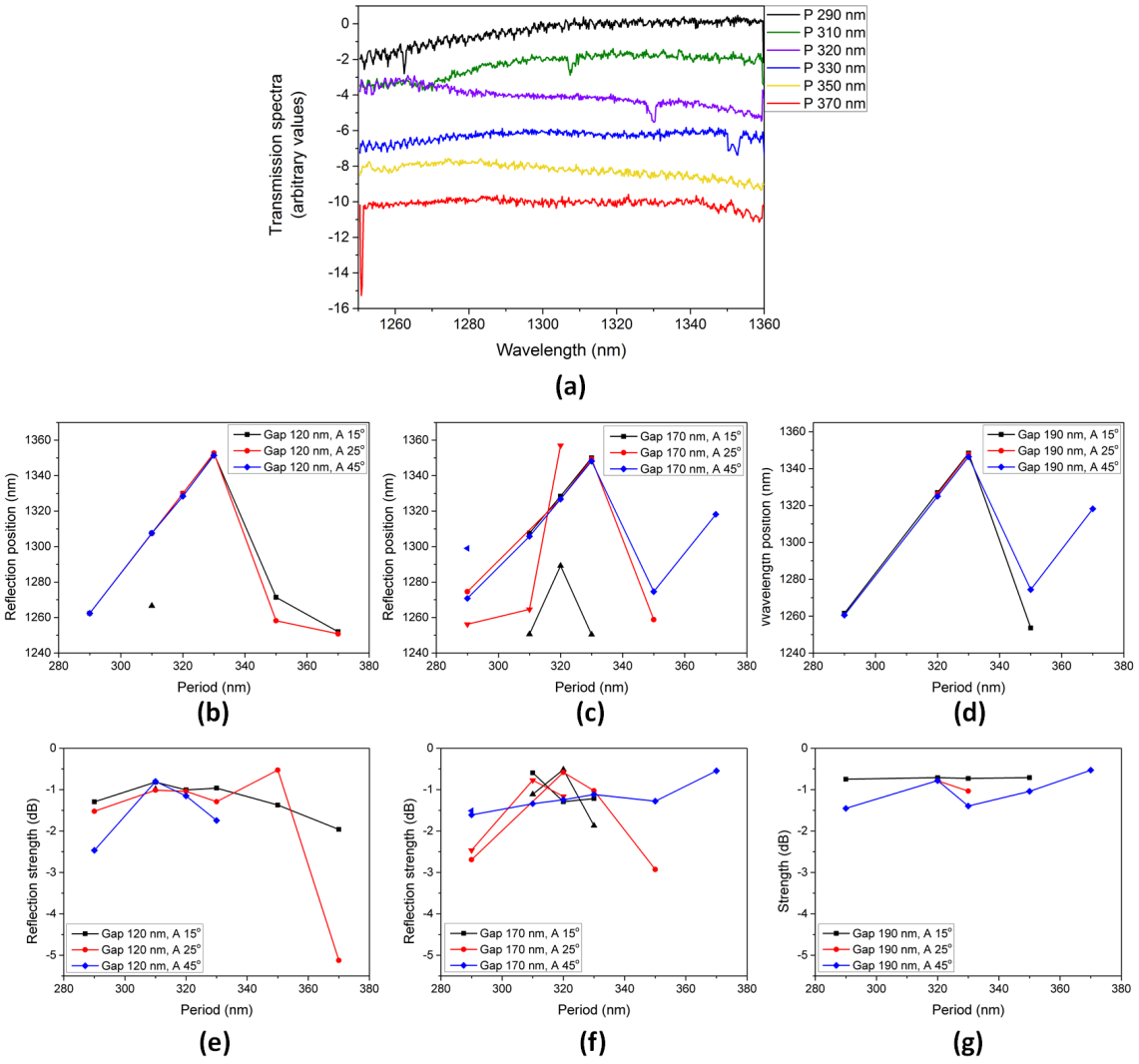


Figure 8.3.7: (a) Example spectra of the continuous width cladding modulated distributed Bragg reflectors with a central waveguide width of $WG = 320\,nm$ and rail width of $Rail = 120\,nm$. The angle is 25° and the gap $G = 120\,nm$. All spectra were normalized to their value at $1300\,nm$ where the stabilization took place before measurement, then for better visualization each spectrum has been spaced with $2\,dB$ compared to the previous. (b-d) Reflection position as a function of the period, datasets for varied gap and angle presented. (e-g) Corresponding dependences of the reflection strength as a function of the period.

There is no similarly clear tendency for the reflection strength to be found as a function of the period. Unlike what was expected from figure 8.3.3(c) no sudden change in strength is recorded when reaching higher wavelengths aside from the single point for $Angle = 25°$ in figure 8.3.7(e). This underlines that the change for the simulation is more likely to be due to the difference in rail width than the change in period. The step-index structures did however exhibit a sudden increase in the strength for the larger periods. Interestingly various combinations of gaps and angles show a single period for which the reflection is significantly decreased, but no evident correlation is found across the various data sets. Examining the reflection strength as a function of the gap shows an overall decrease from smallest to largest gap, however the behaviour at the intermediate value of 170 nm differs for the various angle and period configurations. In both cases an angle of 25° leads to an increase in the grating strength and in both cases where it is 15° it decreases. For the largest angle of 45° the higher period increases while the lower decreases. These tendencies are strong enough that it is unlikely to be due to noise in the

measurement set-up and it underlines how complex the relations between the various design parameters are. This behaviour was alluded to by the simulations shown in figure 8.3.5(e). There is a shift in what gaps exhibit deviances, which is expected to be due to different dimensions of the rail width used in the two analyses. The reflection strength as function of the angle exhibits a behaviour opposite of what would be expected from the simulations, figure 8.3.4(e), where it overall decreases as the angle is increased. Some deviance was expected at lower angles for a number of gaps, but once more in the opposite direction. Line broadening or other fabrication errors may be the cause of this discrepancy, however further investigation would be necessary to verify this.

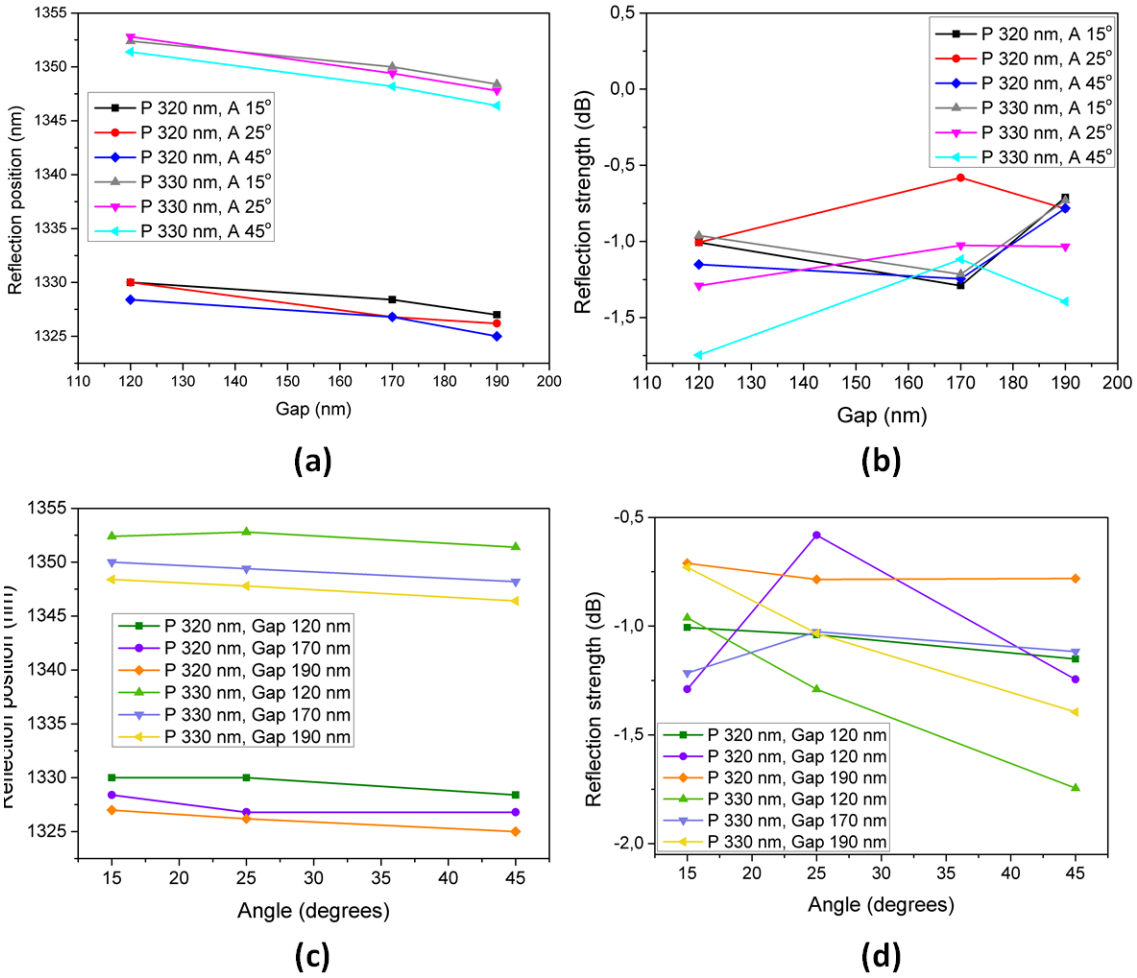


Figure 8.3.8: Correlation between grating design parameters and the reflection behaviour. All graphs include datasets for two different period variations of $P = 320\,nm$ and $P = 330\,nm$. Development of (a) the reflection position and (b) the reflection strength as a function of the gap. Development of (c) the reflection position and (d) the reflection strength as a function of the angle.

As the fabricated devices do not fully match the simulations done ahead of the fabrication a number of additional calculations were made. Figure 8.3.9 shows a comparison of the C-DBRs' behaviour as a function of the design parameters. Figure 8.3.9(a) confirms that these structures exhibit the expected linear increase of the reflection wavelength as a function of the period with the same relation within the range of the TLS. As for the case of the step-index gratings there is a shift in position of the wavelength. The measured reflections do in both cases appear at higher wavelengths with a shift of $\sim 14.5\,\mathrm{nm}$. That the change is so similar across the different chips and designs indicate that it is an artefact of the fabrication, either due to line broadening or a difference in the Si_3N_4 coating layer. The behaviour of the reflection strength is much less in accordance with simulations. It is the be expected that the effect is slightly lower in reality than for the simulation due to degradation, but the overall behaviour would have been expected to to be the same which is only really the case for the angle dependence here. Especially surprising are the points where the measured reflection strength supersedes the simulations as this insinuates shortcomings of the model used or the simulations.

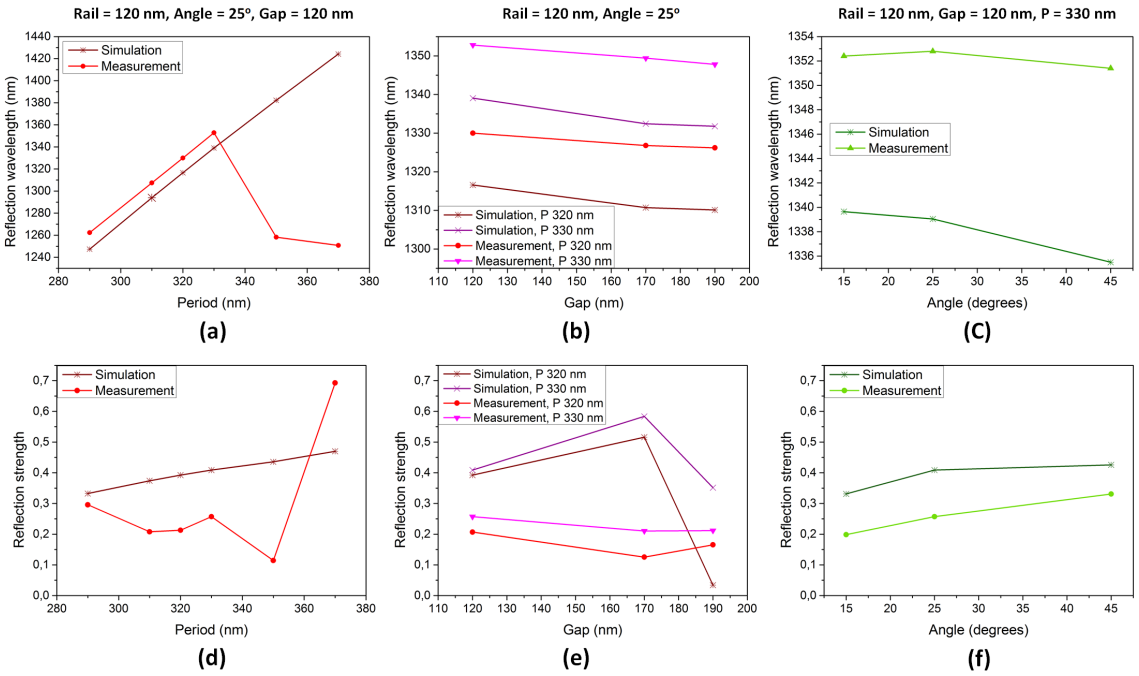


Figure 8.3.9: Direct comparison between simulations and experimental results for a few of the continuous width cladding modulated distributed Bragg reflectors. (a-c) The reflection wavelength as a function of period, gap and angle, respectively. (d-f) The reflection strength as a function of the same three parameters.

Due to the low resolution and the high level of noise in the measurements the development of the bandwidth is difficult to evaluate, however figure 8.3.7(a) may be insinuating an increase as the period increases up until the point where it is only the lower wavelength reflection which is visible.

Figure 8.3.10 shows spectra for a different continuous width C-DBR where the length is $L = 100\,\mu\mathrm{m}$, the central waveguide width is $WG = 320\,\mathrm{nm}$, the rail width $Rail = 150\,\mathrm{nm}$, and the angle is $Angle = 45°$. Three different periods are shown for two different gaps. Two things are especially worth taking note of: Firstly the reflections found for the smaller gap are much stronger than those previously presented, in spite of all previous gratings being six times longer. It may be that of

the remaining device parameters are particularly favourable, as the simulations have shown that even small changes can have a large impact. It may also be due to errors during fabrication as the two data sets were collected from different chips. There is, however, no reason to suspect that longer gratings would have a worse performance if all other parameters are fixed. Other datasets have confirmed the expected relation of grating strength increasing with the length as presented in figure 8.3.1. Secondly it is notable that the change in the gap causes a change in which period exhibits the strongest reflection. To investigate this further simulations were made for these exact grating parameters and the results are shown in figure 8.3.10(c,d). Although the period dependence is as expected for the reflection position, the measured strength differs significantly. That the reflection for the $P = 310\,\text{nm}$ design has practically disappeared is surprising especially for its great strength in the simulations. It is, however, even more curious that the $P = 320\,\text{nm}$ grating is much stronger than expected for the 180 nm gap and given the opportunity it should be investigated whether this behaviour is reproduced on other chips.

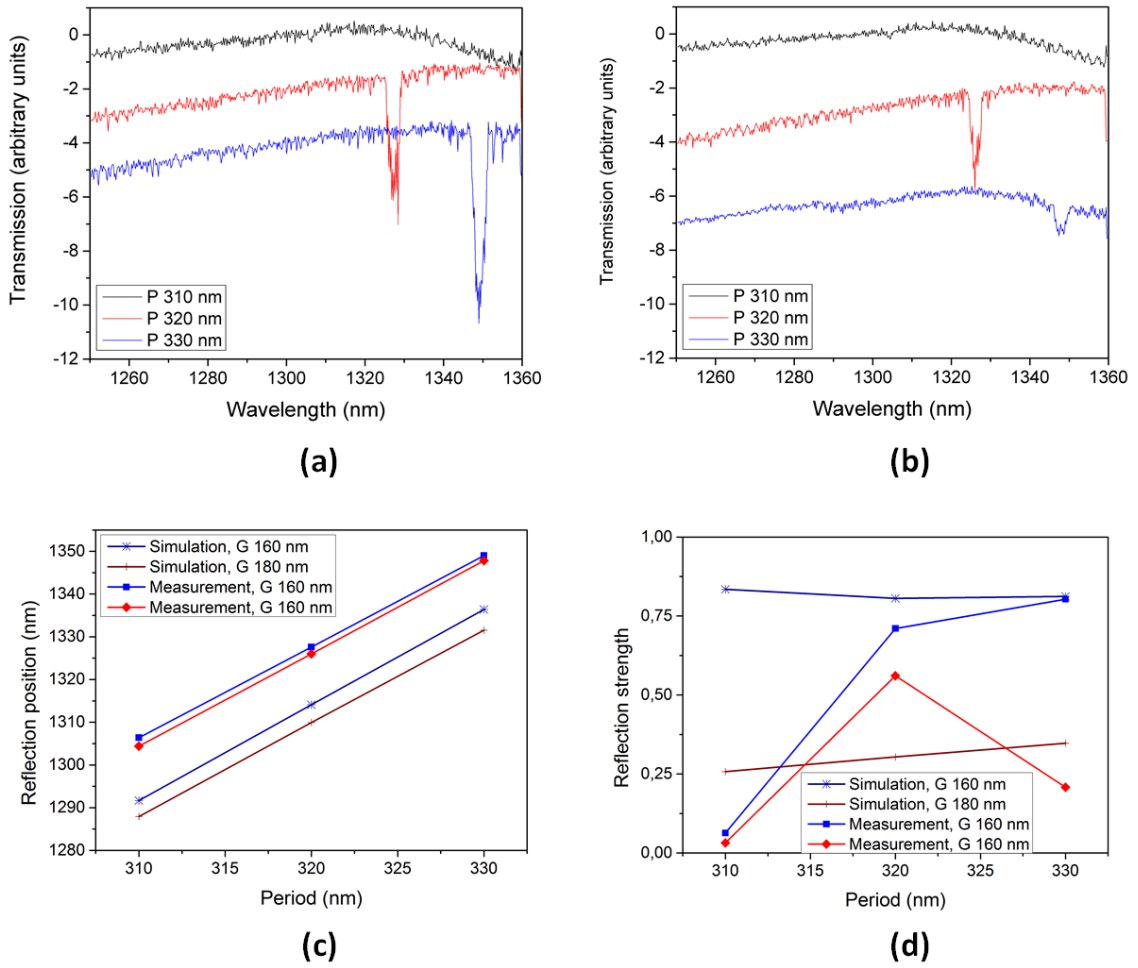


Figure 8.3.10: Spectra of cladding modulated continuous width gratings. Central waveguide width 320 nm, rail width 150 nm, angle 45° for (a) a gap of 160 nm and (b) a gap of 180 nm. The spectra for the various periods are normalized to the value at 1300 nm and subsequently spaced downwards by 2 dB for each spectrum to make hem more easily distinguished. (c,d) Comparison of measured and simulated results correlating the reflection position and strength with the period.

These designs included a variation of the rail width. Figure 8.3.11 reveals that the reflection positions shifts slightly towards higher wavelengths as the rail becomes wider. The trend for the reflection strength depends on period, gap, and rail width alike. For the smaller period the behaviour of the two gap sets are similar while for a period of 330 nm they differ significantly. For a smaller gap of 160 nm the strength increases significantly as the rail width is increased while it decreases, but only slightly, for the larger gap. Figure 8.3.11(c,d) include direct comparison to simulations for the $P = 330\,\text{nm}$ case. The rail dependence of the reflection strength is expected to be linear but are not found for the experiments. Once more it would be necessary to do further measurements on the reproducibility of the measured results to determine whether the discrepancy arises from the model not being adequate or fabrication errors.

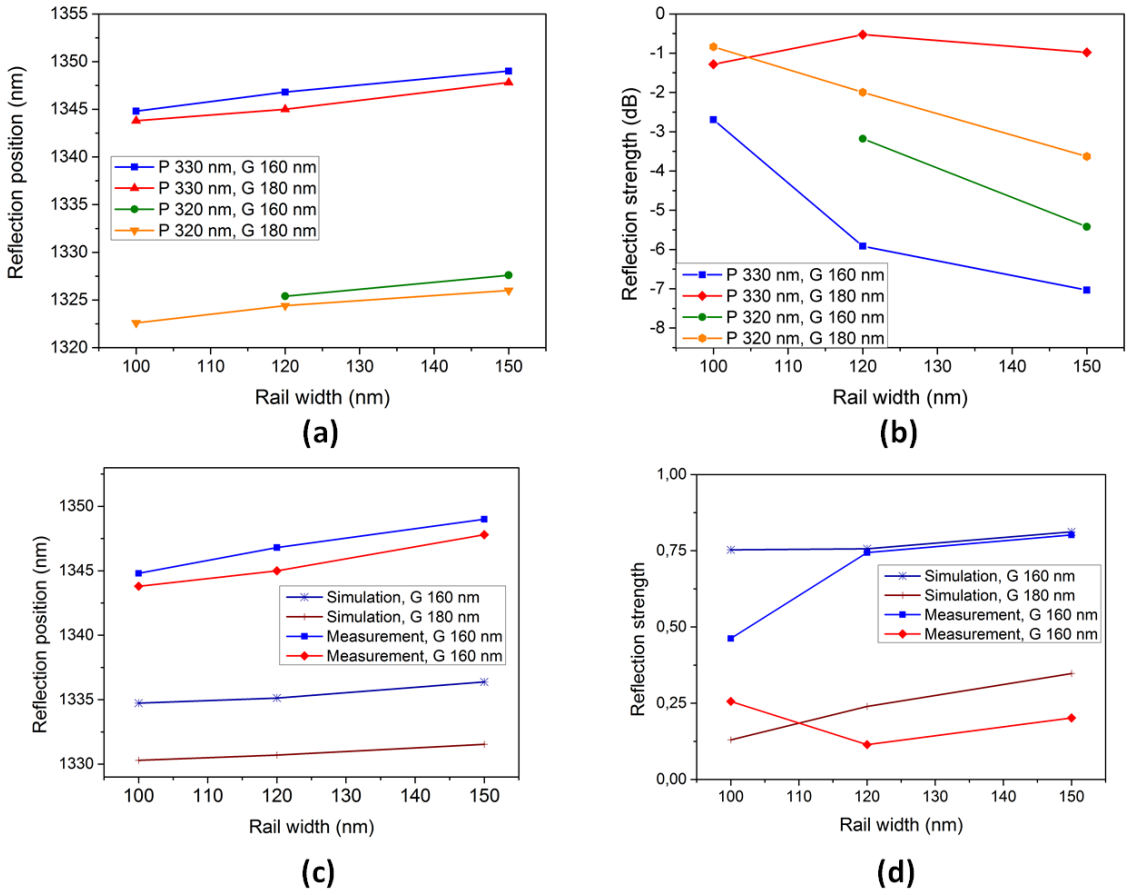


Figure 8.3.11: Development of (a) reflection position and (b) reflection strength as a function of the rail width for a continuous width cladding modulated grating with a central waveguide width of 320 *nm, an arc angle of* $Anle = 45°$*, and a rail width of* 150 *nm. Shown for sets of two different periods and two different gaps. (c,d) Direct comparison of simulations and experimental results for the data-set with a period of* $P = 330\,nm$*.*

The angle dependence is also investigated for these device parameters and the trends are presented in figure 8.3.12. The position of the reflection is only slightly dependent on the angle as would be expected from the simulations. As long as the investigated angles do not deviate far from the peak performance stability is expected. The reflection strength tends to increase as the angle is increased, just as was found for the previous case. The direct comparison between simulations and measurements presented in 8.3.12(c,d) once more shows a nice correspondence for the placement of the reflection, but large deviances for the behaviour of the reflection strength. This may to some extent be explained by the strong noise levels of the set-up and the drift occurring during the long measurement times.

The complex relations between the various parameters will need to be investigated further both experimentally and numerically to obtain a more complete picture.

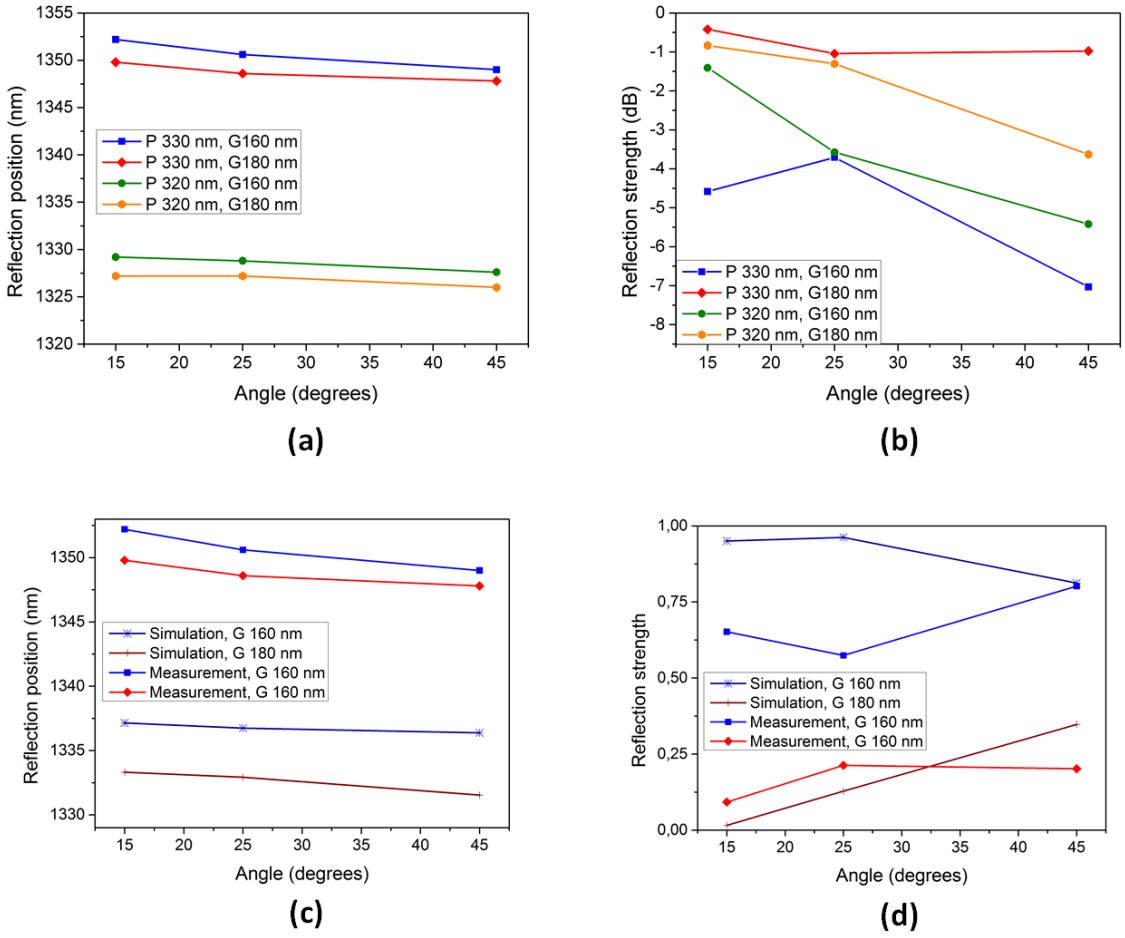


Figure 8.3.12: Development of (a) reflection position and (b) reflection strength as a function of the angle for a continuous width cladding modulated grating with a central waveguide width of 320 nm and a rail width of 150 nm. Results presented for two sets of periods and two different gaps. (c,d) Direct comparison of the simulated and experimental results or the $P = 330$ nm datasets.

Sinusoids

Currently there are no simulations to compare the sinusoid modulated structures to, however they are expected to be similar to the arc gratings. Rather than defining an angle of the bend, the biggest height-difference of the rail is specified and varied throughout. In this manner a large difference in height compares to a large angle. This structure was tested for $dY = 25\,\text{nm}$, $dY = 50\,\text{nm}$, and $dY = 100\,\text{nm}$ height differences. On the chip the period and gap was also varied. All other parameters were fixed so that the central waveguide was $WG = 320\,\text{nm}$, the length of the grating $L = 100\,\mu\text{m}$, and the rail width is 120 nm.

Figure 8.3.13 shows the development of the reflection position as well as the reflection strength as a function of the period for datasets of three different modulation amplitudes and two different gap-distances. It seems that the deviance from the linear relation of the reflection position is no more than a manifestation of the measurement uncertainty and noise. Unfortunately it was not possible to gather the statistics necessary to verify this, however the linear dependence was found for the arc gratings and it is unlikely that it would not carry over to the sinusoid modulation.

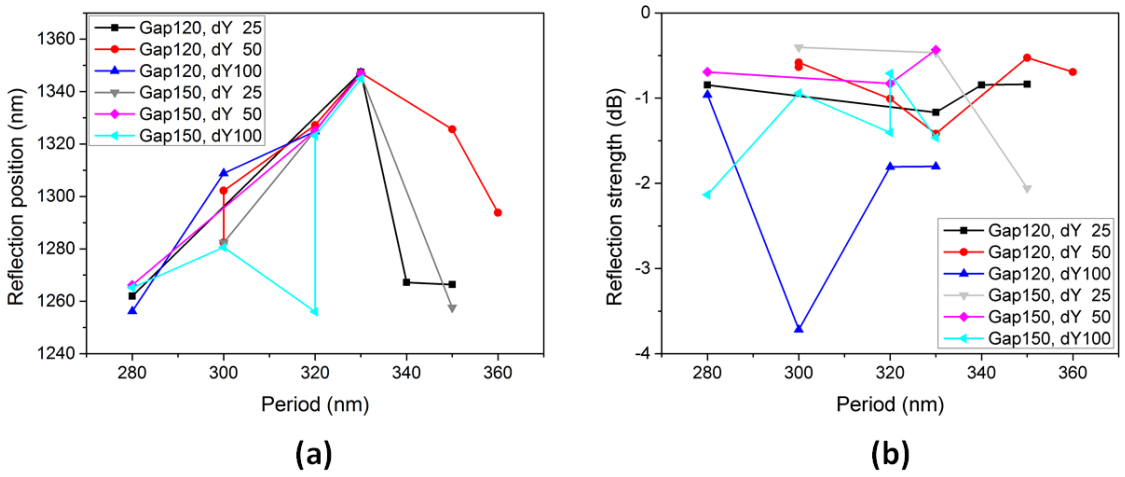


Figure 8.3.13: (a) Reflection position as function of the period for sinusoidal continuous width cladding modulated distributed Bragg reflectors. In both cases shown for the three possible modulation amplitude variations and at two different gap-sizes of 120 nm and 150 nm respectively. (b) Reflection strength as a function of the period for the same data-sets as presented in (a).

Figure 8.3.13(b) reveals no clear trend in the development of the reflection strength of the grating as a function of the period. Aside from the odd outliers it seems that the reflection level is relatively stable across various periods. There is no clear trend as function of the modulation amplitude when comparing the datasets.

In figure 8.3.14 the dependency is instead investigated as a function of the distance between rail and the center waveguide. Comparing various periods and for the different rail modulation amplitudes it is difficult to establish a clear trend as was the case for the simulations of the continuous width structures. For the case of $P = 330\,\text{nm}$ the reflection strength increases as the gap decreases, this holds true for all modulation sizes. This trend does however not carry over to the other cases as illustrated in figure 8.3.14(a) and (b). The $dY = 100\,\text{nm}$ modulation hints that there may be a specific optimum gap distance for each period; in all cases it is however evident that the various modulation strengths are affected in different ways. Generally the biggest sinusoidal modulation of the rail leads to the strongest

reflections, causing them to be detectable for the highest number of periods. It is curious in the example of $P = 280\,\text{nm}$ figure 8.3.14(a), where both of the small dY cases show a tendency of increasing reflection strength as the gap increases yet the reflection completely disappears as the gap increases beyond 120 nm and 170 nm, respectively.

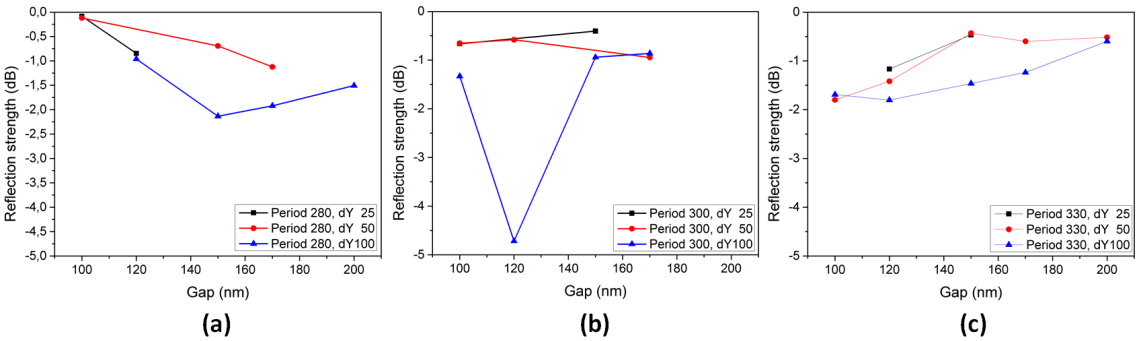


Figure 8.3.14: Depth of the reflection as indicated by the loss that it gave rise to, as a function of the gap between central waveguide and the rail. $WG = 320\,nm$ $Rail = 120\,nm$. The three different amplitudes of the rail modulation is plotted for the three cases of (a) $P = 280\,nm$, (b) $P = 300\,nm$, and (c) $P = 330\,nm$.

The wavelength at which the reflection takes place exhibits no real dependence on the gap. All variations measured were minor and expected to be a result of the poor resolution and potential fabrication errors rather than an actual correlation to device parameters as no trends were found.

8.4 Summary

In this chapter a new type of distributed Bragg reflector has been proposed and demonstrated. Rather than modulating the side walls, the necessary refractive index changes can be introduced through a periodic cladding structure. Here this method has, for the first time, been tested with a continuous structure.

Initially a number of simulations were made using Lumerical ModeSolutions to determine what parameters were suitable for fabrication. Two types of designs were under investigation namely a comb structure with a step-wise index change and a continuous width structure with a sine-like modulation. Numerically both structures showed promise, wherefore they were fabricated for experimental tests. Both types of structures were verified to function as DBRs.

A number of structures were fabricated with various parameter changes allowing for a coarse mapping of how the reflections were affected by design choices. A clear dependence of the reflection wavelength was found on the period however it was, to a smaller degree, influenced by other parameters as well. The effects dictating the strength of the reflection are more complex being highly influenced by all of the design parameters.

It was the hypothesis that the continuous width structures would lead to a better correspondence between simulations and fabricated structures than would the step index structures. This was qualitatively verified by the SEM images of the devices as the sharp corners of the step index grating exhibited noticeable rounding. Furthermore some repeated deviance was found for the relation between the period and reflection position for the step index, but not for the continuous width case. This strengthens the hypothesis, but as other relationships did not match the simulations in either of the cases it cannot be deemed verified.

The work carried out in this project has demonstrated the feasibility of creating strong DBRs, reflecting up to 80 % of the signal, through continuous cladding modulation. Further investigation will have to be conducted to gain proper knowledge of how to engineer the reflection spectra and obtain the desired grating effects. Firstly it would be necessary to do more measurement to collect statistics on the grating performance in an effort to correlate it to the simulations. As it currently seems that there are some deviations, the first step will be to find the root of these discrepancies. Then it would be possible to tune the model to perform a further numerical study. Following that another experimental investigation should be undertaken to verify the second iteration of the model. A key in this correlation between theoretical and experimental results is also to refine the experiential set-up to eliminate the uncertainties introduced by the high noise floor, the coarse wavelength scan and the drift of the system.

CHAPTER 9

FUTURE PROSPECTS AND CONCLUSION

As the interest for optical communication increases, much effort is being put into creating components that may together form the basis of photonic integrated circuits. The main focus of this project has been to investigate the potential of the inverse design method topology optimization. Topology optimization is an especially useful tool when dealing with the challenge of fully exploiting the chip area.

A number of different nanophotonic components have been designed, developed and experimentally verified.

Mode division multiplexing is a suggested means of increasing the transfer capacity. The key devices of mode converters and multiplexers have been designed during this project. The height difference of the in- and out-put waveguides was found to be an important feature for these structures. Although it was exploited that TO can be used on an arbitrary starting point structure and still obtain functional devices, it became clear that choosing a good starting point structure is essential for optimum functionality. Finding the good starting point can potentially be done by manually iterating on the topology optimized designs as was demonstrated for the compact TE_0-to-TE_1 mode converter. Furthermore once a well-functioning device has been obtained for the lowest possible complexity design, this can be used as the off-set for devices with increased functionality. This was tested, for example, for the 3-mode multiplexer. A strength of TO is how increasing the complexity in this manner requires a relatively little increase in footprint as the functionality can be maintained in a single step rather that requiring cascading of multiple devices.

A challenge of the method of TO is that the complex scattering functionalities of the devices rely on small feature sizes to be able to realize the small footprints. Using a wavelength multiplexer as the basis for the investigation, the relation between size and performance was examined. It was found that for a particular design configuration there would be device dimensions most likely to converge to a good local optimum. This means that, although it can be done with tweaking, it will not always lead to lower losses to increase the size of the device. However, it does require that individual analyses are made for any device under investigation to obtain an optimum configuration. Less encouraging were the results concerning the feature sizes within the design. It was found that smaller feature sizes would generally lead to better performance as it gives the TO more freedom. Meanwhile employing filters in multiple steps, applying first large filters and subsequently smaller ones,

may help alleviate this problem and create low-loss structures with larger features. This is important to obtain robustness as the small features of the designs make them vulnerable to fabrication errors.

The final design type investigated with TO was the compact taper connecting fundamental modes of PhWs of various widths. Unfortunately the challenge of stably exciting the fundamental mode in wide waveguides barred the realization of these devices on the desired scale. The method was however demonstrated to have potential could this challenge be overcome as it was applicable for smaller dimensions.

Topology optimization has great potential for the designs of many more devices. Using inverse design methods gives the designers the challenge of coming up with interesting new functionalities rather than how to realize them. The next steps might however be to apply it to existing applications in an attempt to get better functionality and smaller footprints. This could for example be the design of high-q cavities, crossing interconnects, antennas and fibre couplers. It is however not a method without limitations. There are many functionalities that could and should be implemented in the software to obtain even better devices.

The work with the compact taper, presented in chapter 7, showed that the implementation of a mode exciter and/or a mode objective is necessary for some devices. Furthermore it would make other devices, such as the modemultiplexers of chapter 5, more straight forward to design. For the design of multiplexers handling many modes it might even be a necessity. Another issue that the tapers revealed was that of symmetry. Even though strange and unexpected structures are a forte of TO, sometimes it is not necessarily the best solution. For symmetric mode expansion as in this case, it is safe to assume that the optimum design is symmetric and by enforcing the symmetry it is possible to get closer to the global optimum. In addition enforcing symmetry can also lower the calculation time of the optimization.

The issue of convergence towards local optima is one of the general challenges of inverse design methods. Although it is not a way of ensuring that the global optimum is reached, it may be possible to obtain better optima by the use of refined - or even random - starting points. As shown in figure 2.1.2 the overall features of the design are reached quickly within the first 15 iterations or so. Allowing this initial part of the optimization to be redone for multiple starting point configurations and then only finishing the full optimization of a few devices that exhibit the best objective values after the first iterations may guide structures towards better optima. Ideally the Si in the DD could be segmented and randomly or systematically distributed within these to generate relevant starting point structures. A more refined version of this could be implemented using a large filter size for the first 15 iterations. Using the large filter's initial result as a starting point, a full length optimization with a more appropriate filter size can be performed to obtain better results, as touched upon in section 6.3.2.

Incorporating the use of more materials in the optimization has already received some amount of interest [75]. This capability is definitely worth incorporating into the software as well. For work with plasmonics, especially, the incorporation of metals is a necessity but there are many more relevant applications.

While one of the great benefits of TO is that it can produce compact structures with small footprints, this requires that the feature sizes are very small too. Although the devices are readily fabricated using e-beam lithography, that will not be feasible on large scale as needed for commercialization. For future work with TO it is thus necessary to address the trade-off between size and functionality. Especially

the usage and implementation of refined filtering mechanisms is important for this perspective.

Similarly the overall robustness to fabrication imperfections is an important issue and ideally robustness ensuring features should be incorporated into the software. Only lightly touched upon in this thesis, there is also the issue of heat compatibility. Including experimental investigation of the robustness to heat change would be interesting along with exploring the potential of designing structures meant for operation under elevated temperatures.

Finally a novel method of creating distributed Bragg reflectors was proposed and demonstrated. The design of these did not include TO but a more Edisonian approach. The structures are modulated by a pair of symmetric continuous rails placed next to the central waveguide in which the light is injected. The effect of varying the design parameters was investigated through both simulation and measurements. The relations are complex and further investigation will be necessary to build a model of the effect. It has however been demonstrated that these types of structures can be used to create reflector gratings.

This project has thus contributed new devices for the toolbox needed for photonic integrated circuits. It has also furthered the understanding of the method of TO employed for nanophotonic devices. There is still a huge potential for the utilization of this tool. Although there are a number of ways in which the method can still be strengthened, its universality suggests that the main challenge of the future design of passive nanophotonic components will be coming up with the structures to use inverse design for.

REFERENCES

[1] "White paper: Cisco VNI forecast and methodology." `http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.html`, 2016.

[2] "What is big data?." `https://www-01.ibm.com/software/data/bigdata/what-is-big-data.html`, 2009.

[3] D. J. Richardson, J. M. Fini, and L. E. Nelson, "Space-division multiplexing in optical fibers," *Nature Photonics*, vol. 7, 2013.

[4] M. J. Paniccia, "A perfect marriage: optics and silicon," *Optik & Photonik*, vol. 6, 2011.

[5] R. Kirchain and L. Kimerling, "A roadmap for nanophotonics," *Natur photonics*, vol. 1, 2007.

[6] B. Jalali and S. Fathpour, "Silicon photonics," *Journal of Lightwave technology*, vol. 24, 2006.

[7] G. E. Jellison jr. and H. H. Burke, "The temperature dependence of the refractive index of silicon at elevated temperatures at several laser wavelengths," *Journal of Applied Physics*, vol. 60, 1986.

[8] M. P. Bendsøe and N. Kikuchi, "Generating optimal topologies in structural design using a homogenization mehod," *Computer Methods in Applied Mechanics and Engineering*, vol. 71, 1988.

[9] "DTU topology optimization website." `http://www.topopt.dtu.dk/`, 2009.

[10] O. Sigmund, "Design of multiphysics actuators using topology optimization - part 1," *Computer Methods in Applied Mechanics and Engineering*, vol. 190, 2001.

[11] C. S. Andreasen, A. R. Gersborg, and O. Sigmund, "Topology optimization of microfluidic mixers," *International Journal for Numerical Methods in Fluids*, vol. 61, pp. 498–513, 2009.

[12] P. I. Borel, A. Harpøth, L. H. Frandsen, M. Kristensen, P. Shi, J. S. Jensen, and O. Sigmund, "Topology optimization and fabrication of photonic crystal structures," *Optics Express*, vol. 12, 2004.

[13] J. S. Jensen and O. Sigmund, "Topology optimization for nano-photonics," *Laser & Photonics Reviews*, vol. 5, 2011.

[14] Y. Elesin, B. S. Lazarov, J. S. Jensen, and O. Sigmund, "Design of robust and efficient photonic switches using topology optimization," *Photonics and Nanostructures*, vol. 10, 2012.

[15] Y. Elesin, B. S. Lazarov, J. S. Jensen, and O. Sigmund, "Time domain topology optimization of 3d nanophotonic devices," *Photonics and Nanostructures - Fundamentals and Applications*, vol. 12, 2014.

[16] L. Li and K. Khandelwal, "Volume preserving protection filters and continuation methods in topology optimization," *Engineering strctures*, vol. 85, 2015.

[17] Y. Jia, S. F. Mingaleev, M. Schillinger, D. A. B. Miller, S. Fan, and K. Busch, "Wannier basis design and optimization of photonic crystal waveguide crossing,"

[18] J. Lu, C. Petre, and E. Yablonovitch, "Numerical optimization of a grating coupler for the efficient excitatio nof surface plasmons at an Ag-SiO_2 interface,"

[19] P. Sanchis, P. Villalba, F. Cuesta, A. Håkansson, A. Griol, J. V. Galán, A. Brimont, and J. Martí, "Highly efficient crossing structure for silicon-on-insulator waveguides," *Optics Letters*, vol. 34, pp. 2760–2762, 2009.

[20] J. Lu and J. Vuckovic, "Objective-first design of high-efficiency, small-footprint couplers between arbitrary nanophotonic waveguide modes," *Optics Express*, vol. 20, pp. 7221–7236, 2012.

[21] J. Lu and J. Vuckovic, "Nanophotonic computational design," *Optics Express*, vol. 21, pp. 13351–13367, 2013.

[22] A. Y. Piggot, J. Lu, K. G. Lagoudakis, J. Petyiekwicz, T. M. Babinec, and J. Vuckovic, "Inverse design and demonstration of a compact and broadband on-chip wavelength demultiplexer," *Nature Photonics*, vol. 9, 2015.

[23] B. Shen, P. Wang, R. Polson, and R. Menon, "An integrated-nanophotonics polarization beamsplitter with $2.4x2.4m^2$ footprint," *Nature photonics*, vol. 9, pp. 378–382, 2015.

[24] J. C. C. Mak, C. Sideris, J. Jeong, A. Hajimiri, and J. K. S. Poon, "Binary particle swarm optimized 2 x 2 power splitters in a standard foundry silicon photonic platform," *Optics Letters*, vol. 41, pp. 3868–3871, 2016.

[25] J. kennedy, "The particle swarm: Social adaptation of knowledge," *IEEE International Conference on Evolutionary Computation*, pp. 303–308, 1997.

[26] O. Sigmund, J. S. Jensen, and L. H. Frandsen, "On nanostructured silicon sucess," *Nature photonics*, vol. 10, pp. 142–143, 2016.

[27] E. Khomtchenko, "Manual for duv stepper." `http://labmanager.dtu.dk/d4Show.php?id=2519&mach=273`, 2013.

[28] M. Pu, L. Liu, H. Ou, K. Yvind, and J. M. Hvam, "Ultra-low-loss inverted taper coupler for silicon-on-insulator ridge waveguide," *Optics Communications*, vol. 283, 2010.

[29] T. Greibe, "User manual for e-beam writer jbx-9500." `http://labmanager.dtu.dk/d4Show.php?id=3566&mach=292`, 2015.

[30] L. H. Frandsen, *Fabrication and Characterization of Photonic Bandgap Components*. PhD thesis, Technical University of Denmark, 2006.

[31] J. M. Lindhard, "Manual for ase." `http://labmanager.dtu.dk/d4Show.php?id=1609&mach=105`, 2015.

[32] Y. Ding, H. Ou, J. Xu, M. Xiong, and C. Peucheret, "On-chip mode multiplexing based on a single grating coupler," *IEEE Photonics Conference*, 2012.

[33] Y. Huang, G. Xu, and S.-T. Ho, "An ultracompact optical mode order converter," *IEEE Photonics Thechnology Letters*, vol. 18, 2006.

[34] G. Chen and J. U. Kang, "Waveguide mode converter based on two-dimensional photonic crystals," *Optics Letters*, vol. 30, 2005.

[35] V. Liu, D. A. B. Miller, and S. Fan, "Ultra-compact photonic crystal waveguide spatial mode converter and its connection to the optical diode effect," *Optics Express*, vol. 20, 2012.

[36] L. H. Frandsen, A. Harpøth, P. I. Borel, M. Kristensen, J. S. Jensen, and O. Sigmund, "Broadband photonic crystal waveguide 60° bend obtained utilizing topology optimization," *Optics Express*, vol. 12, 2004.

[37] P. I. Borel, B. Bilenberg, L. H. Frandsen, T. Nielsen, J. Fage-Pedersen, A. V. Lavrinenko, J. S. Jensen, O. Sigmund, and A. Kristensen, "Imprinted silicon-based nanophotonics," *Optics Express*, vol. 15, 2007.

[38] S. Boscolo, M. Midria, and T. F. Krauss, "Y junction in phtonic crystal channel waveguides: High transmission and impedance matching," *Optics Letters*, vol. 27, 2002.

[39] L. H. Frandsen, Y. Elesin, L. F. Frellsen, M. Mitrovic, Y. Ding, O. Sigmund, and K. Yvind, "Topololgy optimized mode conversion in a waveguide fabricated in silicon-on-insulator material," *Optics Express*, vol. 22, 2014.

[40] Q. Zhang, C.-W. Yan, and L. Liu, "Theoretical design and analysis for te20-te10 rectangular waveguide mode converters," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, 2012.

[41] H. Guan, Y. Ma, R. Shi, A. Novack, J. Tao, Q. Fang, A. E.-J. Lim, G.-Q. Lo, T. Baehr-Jones, and M. Hochberg, "Ultracompact silicon-on-insulator polarization rotator for polarization-diversified circuits," *Optics Letters*, vol. 39, 2014.

[42] Y. Ding, J. Xu, F. da Ros, B. Uan, H. Ou, and C. Peucheret, "On-chip two-mode division multiplexing using tapered direction coupler-based mode multiplexer and demultiplexer," *Optics Express*, vol. 21, 2013.

[43] J. Wang, S. He, and D. Dai, "On-chip silicon 8-channel hybrid (de)multiplexer enabling simultaneous mode- and polarization-division-multiplexing," *Laser & Photonics Reviews*, 2014.

[44] W. Chen, P. Wang, Y. Zhang, and J. Yang, "Design of an ultra-broadband silicon mode (de)multiplexer," *Optics Communications*, vol. 363, 2015.

[45] J. D. Love and N. Riesen, "Single-, few- and multimode y-junctions," *Journal of Lightwave Technology*, vol. 30, 2012.

[46] J. B. Driscoll, R. R. Grote, B. Souhan, J. I. Dadap, M. Lu, and R. M. Osgood jr., "Asymmetric y junction in silicon waveguide for on-chip mode-division multiplexing," *Optics Letters*, vol. 38, 2013.

[47] K. Shirafuji and S. Kurazono, "Transmission characteristics of optical asymmetric y junction with a gap region," *Journal of Lgihtwave technology*, vol. 9, 1991.

[48] Y. Kawaguchi and K. Tsutsumi, "Mode multiplexing and demultiplexing devices using multimode interference couplers," *Electronics Letters*, vol. 38, 2002.

[49] T. Uematsu, Y. Ishizaka, Y. Kawaguchi, K. Saitoh, and M. Koshiba, "Design of a compact two-mode multi/demultiplexer consisting of multimode interference waveguides and a wavelength-insensitive phase shifter for mode-division multiplexing transmission," *Journal of Lightwave Technology*, vol. 30, 2012.

[50] D. Pérez-Galacho, D. Marris-Morini, A. Ortega-Moñux, J. G. Wangüemert-Pérez, and L. Vivien, "Add/drop mode-division multiplexer based on a mach-zehnder interferometer and periodic waveguides," *IEEE Photonics Journal*, vol. 7, 2015.

[51] L.-W. Lou, N. Ophir, C. P. Chen, L. H. Gabrielli, C. B. Poitras, K. Bergmen, and M. Lipson, "Wdm-compatible mode-division multiplexing on a silicon chip," *Nature Communications*, 2013.

[52] B. A. Dorin and W. N. Ye, "Two-mode division multiplexing in a silicon-on-insulator ring resonator," *Optics Express*, vol. 22, 2014.

[53] G. J. Veldhuis, J. H. Berends, and P. V. Lambeek, "Design and characterization of a mode-splitting ψ -junction," *Journal of Lightwave Technology*, vol. 14, 1996.

[54] "Wavelength division multiplexing (WDM)." `http://www.thefoa.org/tech/dwdm.htm`, 2013.

[55] M. S. Dahlem, C. W. Holzwarth, A. Khilo, F. X. Kärtner, H. I. Smith, and E. P. Ippen, "Reconfigurable multi-channel second-order silicon microring-resonator filterbanks for on-chip wdm systems," *Optics Express*, vol. 19, 2011.

[56] F. Xia, M. Rooks, L. Sekaric, and Y. Vlasov, "Ultra-compact high order ring resonator filters using submircon silicon photonic wires for on-chip optical interconnects," *Optics Express*, vol. 15, 2007.

[57] F. Horst, W. M. J. Green, B. H. Offrein, and Y. A. Vlasov, "Silicon-on-insulator echelle grating WDM demultiplexers with two stigmatic points," *IEEE Photonics Technology Letters*, vol. 21, 2009.

[58] K. Yamada, T. Shoji, T. Tsuchizawa, T. Watanabe, J.-I. Takahashi, and S.-I. Itabashi, "Silicon-wire-based ultrasmall lattice filters with wide free spectral ranges," *Optics Letters*, vol. 28, 2003.

[59] T. Fukazawa, F. Ohno, and T. Baba, "Very compact arrayed-waveguide-grating demultiplexer using Si photonic wire waveguides," *Japanese Journal of Applied Physics*, vol. 43, 2004.

[60] S. Chen, Y. Shi, S. He, and D. Dai, "Compact monolithically-integrated hybrid (de)multiplexer based on silicon-on-insulator nanowires for PDM-WDM systems," *Optics Express*, vol. 23, 2015.

[61] Y. Fu, T. Ye, W. Tang, and T. Chu, "Efficient adiabatic silicon-on-insulator waveguide taper," *Photonics Research*, vol. 2, 2014.

[62] T. Ye, Y. Fu, L. Qiao, and T. Chu, "Low-crosstalk Si arrayed waveguide grating with parabolic tapers," *Optics Express*, vol. 22, no. 26, pp. 31899–31906, 2014.

[63] B. Luyssaret, P. Bienstman, and R. Baets, "A versatile spot-size converter design," *Proceedings of European Conference on Optical Communication*, 2004.

[64] D. Vermeulen, K. V. Acoleyen, S. Ghosh, S. Selvaraja, W. A. D. de Cort, N. A. Yeob, E. Hallynck, K. de Vos, P. P. P. Debackere, P. Dumon, W. Bogaerts, G. Roelkens, D. W. Thourhout, and R. Bbaets, "Efficient tapering to the fundamental quasi-tm mode in asymmetrical waveguides," *Proceedings of European Conference on Integrated Optics*, 2010.

[65] K. van Acoleyen and R. Baets, "Compact lens-assisted focusing tapers fabricated on silicon-on-insulator," *Proceedings of IEEE Converence on Group IV Photonics*, 2011.

[66] K. O. Hill and G. Meltz, "Fiber Bragg grating technology fundamentals and overview," *Journal of Lightwave Technology*, vol. 15, 1997.

[67] T. Coroy, *Wavelength Measurement System for Bragg Fiber Optic Sensors based on Quantum Well Electroabsorption Photodetectors*. PhD thesis, University of Toronto Institute for Aerospace Studies, 1999.

[68] T. Erdogan, "Fiber grating spectra," *Journal of Lightwave Technology*, vol. 15, 1997.

[69] J. H. Schmid, P. Cheben, P. J. Bock, R. Halir, J. Lapointe, S. Janz, A. Delâge, A. Densmore, J.-M. Fédeli, T. J. Hall, B. Lamontagne, R. Ma, I. Molina-Fernández, and D.-X. Xu, "Refractive index engineering with subwaelength gratings in silicon microphotonic waveguides," *IEEE Photonics Journal*, vol. 3, 2011.

[70] D. C. Flanders, H. Kogelnik, R. V. Schmidt, and C. V. Shank, "Grating filters for thin-film optical waveguides," *Applied Physics Letters*, vol. 24, 1974.

[71] H.-C. Kim, K. Ikeda, and Y. Fainman, "Resonant waveguide devices with vertical gratings," *Optics Letters*, vol. 32, 2007.

[72] O. Bondarenko, Q. Gu, J. Shane, A. Simic, B. Slutsky, and Y. Fainman, "Wafer bonded distributed feedback laser with sidewall modulatd bragg gratings," *Applied physics letters*, vol. 103, 2013.

[73] D. T. H. Tan, K. Ikeda, and Y. Fainman, "Cladding-modulated bragg gratings in silicon waveguides," *Optics Letters*, vol. 34, 2009.

[74] "Eigenmode exmpansion (EME) solver." `https://www.lumerical.com/tcad-products/mode/EME`.

[75] J. Andkjær, S. Nishiwaki, T. Nomura, and O. Sigmund, "Topology optimization of grating couplers for the efficient excitation of surface plasmons," *J. Opt. Soc. Am. B*, vol. 27, 2009.

[76] I. Giuntoni, A. Gjada, M. Krause, R. Steingrüber, J. Bruns, and K. Petermann, "Tunable Bragg reflectors on silicon-on-insulator rib waveguides," *Optics Express*, vol. 17, 2009.